

研究生教材

YANJIUSHENGJIAOCAI

应用随机过程

YINGYONGSUIJIGUOCHENG

主编　陈家清　赵华玲　梅顺治



武汉理工大学出版社

WUTP Wuhan University of Technology Press



武汉理工大学创新人才培养项目资助

项目策划：理工事业部
责任编辑：彭佳佳
装帧设计：芳华时代
电　话：87515798
http: //www. techbook. com.cn
E-mail:chenjd@whut.edu.cn

武汉理工大学出版社发行部
地　址：武汉市洪山区珞狮路122号
电　话：（027）87523148
传　真：（027）87165708

ISBN 978-7-5629-4193-4



定价：28.00元

应用随机过程

主　编　陈家清　赵华玲
梅顺治

武汉理工大学出版社
·武　汉·

内 容 提 要

本书主要内容包括:概率论基础知识、随机过程的概念和基本类型、泊松过程、离散状态和连续时间马尔可夫链、鞅和布朗运动、随机分析等。

本书尽可能地简化复杂的抽象证明或推导,主要突出知识的应用教学与学习;叙述通俗、简明且例题较多,并给出了章节习题的大部分参考答案,便于读者自学参考。

本书可作为高等院校工科类以及管理、经济与金融类研究生和高年级本科生的教材,同时也可作为广大从事相关专业技术研究人员的参考或自学书籍。

图书在版编目(CIP)数据

应用随机过程/陈家清,赵华玲,梅顺治主编. —武汉:武汉理工大学出版社,2014.9

ISBN 978-7-5629-4193-4

Ⅰ.①应… Ⅱ.①陈… ②赵… ③梅… Ⅲ.①随机过程 Ⅳ.①O211.6

中国版本图书馆 CIP 数据核字(2014)第 145955 号

项目负责人:陈军东　彭佳佳　　**责任编辑**:彭佳佳
责 任 校 对:王　思　　**装帧设计**:芳华时代
出 版 发 行:武汉理工大学出版社
社　　址:武汉市洪山区珞狮路 122 号
邮　　编:430070
经　　销:各地新华书店
印　　刷:安陆市鼎鑫印务有限责任公司
开　　本:787×960　1/16
印　　张:12.5
字　　数:270 千字
版　　次:2014 年 9 月第 1 版
印　　次:2014 年 9 月第 1 次印刷
印　　数:1—2000 册
定　　价:28.00 元
凡购本书,如有缺页、倒页、脱页等印装质量问题,请向出版社发行部调换。
本社购书热线电话:027-87523148　87664138　87515798　87165708(传真)

前　　言

随机过程理论是现代概率论中的一个重要分支,在生物、物理、系统工程、管理和经济等领域都有着广泛的应用。目前国内大多数高校大都在工科、管理和经济等相关专业开设了应用随机过程课程。本书是编者结合多年的教学经验,在自编讲义的基础上,根据教学课时的安排及工科等专业研究生掌握概率统计专业基础知识的程度,以及全国工科院校硕士研究生应用随机过程课程的教学基本要求而编写的。

本书在内容的选取方面,既涵盖了随机过程基本内容、基本思想和基本方法,又通过典型实例的分析,重点突出了知识的实际应用背景,旨在培养学生应用随机过程的理论与方法解决实际问题的能力。同时,尽可能地简化某些结论的抽象证明,突出主要知识的讲解和应用,以期满足学生掌握运用知识的需要。本书大致分为五个部分:概率论基础知识以及随机过程的基本内容(第 1 章、第 2 章),泊松过程(第 3 章),马尔可夫过程(第 4 章、第 5 章),鞅和布朗运动(第 6 章),随机分析(第 7 章)。

参加本书编写的人员有陈家清、梅顺治和赵华玲等老师,他们多年来从事研究生、本科生随机过程的教学工作,具有较为丰富的教学实践经验。同时感谢周树民教授对本书书稿的编写给出的详细指导和修改。

这里特别感谢武汉理工大学研究生院培养处的各位同仁,他们帮助和鼓励编者完成了这本书的编写。余旌胡教授认真审阅了书稿,并提出了许多宝贵的意见,在此表示由衷的感谢!同时还要感谢武汉理工大学出版社对本书出版提供的支持和帮助!

由于编者水平有限,书中的缺点在所难免,敬请读者批评指正。

编　者

2014 年 7 月于马房山

目　　录

第1章　概率论基础知识

1.1　随机事件及其概率

随机试验是概率论的一个基本概念.一个试验(或观察),若它的结果预先无法确定,但具有如下三个性质,则称之为**随机试验**,简称为试验:

(1) 试验的可重复性,即在相同的条件下可重复进行;

(2) 试验的可能结果不止一个,且所有可能结果是已知的;

(3) 试验结果的不确定性,即每次试验前不能确定哪个结果会出现.

随机试验的可能结果称为样本点或基本事件,记为 ω. 所有样本点的集合称为样本空间,记作 Ω. 随机事件(简称为事件)是样本空间 Ω 的子集,包含若干个样本点,用大写字母 A,B,C 表示. 样本空间 Ω 称为必然事件,空集 $\varnothing$ 称为不可能事件. 如果 A 所包含的样本点中有一个发生,则称事件 A 发生. 由 Ω 中的若干子集构成的集合称为集族,用花写字母 $\mathscr{A},\mathscr{B},\mathscr{F}$ 等表示.

一般并不把 Ω 的一切子集都作为事件,因为这将给定义概率带来困难;另一方面,又必须把问题中感兴趣的事件都包括进来,例如 A 是事件,则应要求 $\overline{A}$ 也是事件;若 A,B 是事件,则 $A\cup B,AB$ 也应是事件. 为此,引入 σ- 代数的概念:

定义 1.1　设随机试验 E 的样本空间为 Ω,$\mathscr{F}$ 是 Ω 的子集组成的集族,满足:

(1) $\Omega\in\mathscr{F}$;

(2) 若 $A\in\mathscr{F}$,则 $\overline{A}\in\mathscr{F}$;

(3) 若 $A_i\in\mathscr{F}\ (i=1,2,\cdots)$,则 $\bigcup\limits_{i=1}^{+\infty}A_i\in\mathscr{F}$;

则称 $\mathscr{F}$ 为 Ω 的一个 **σ- 代数**(σ- 域),$\mathscr{F}$ 中的集合称为事件,$(\Omega,\mathscr{F})$ 称为**可测空间**.

可测空间具有如下性质:

(1) $\varnothing\in\mathscr{F}$;

(2) 若 $A_i\in\mathscr{F}(i=1,2,\cdots)$,则有 $\bigcap\limits_{i=1}^{+\infty}A_i\in\mathscr{F}$;

(3) 若 $A_i\in\mathscr{F}(i=1,2,\cdots,n)$,则 $\bigcup\limits_{i=1}^{n}A_i\in\mathscr{F}$,或 $\bigcap\limits_{i=1}^{n}A_i\in\mathscr{F}$;

(4) 若 $A\in\mathscr{F},B\in\mathscr{F}$,则 $A-B\in\mathscr{F}$.

定义 1.2 设$(\Omega,\mathscr{F})$是一可测空间,定义在$\mathscr{F}$上的实值集函数$P(\cdot)$,若满足:

(1) 非负性:对$\forall A\in\mathscr{F}, 0\leqslant P(A)\leqslant 1$;

(2) 规一性:$P(\Omega)=1$;

(3) 完全可加性:对$\forall A_i\in\mathscr{F}(i=1,2,\cdots), A_i\cap A_j=\varnothing, i\neq j$,有

$$P(\bigcup_{i=1}^{+\infty}A_i)=\sum_{i=1}^{+\infty}P(A_i)$$

则称P是$(\Omega,\mathscr{F})$上的概率测度,简称为概率,$(\Omega,\mathscr{F},P)$称为概率空间,$P(A)$表示事件A的概率.

由定义 1.2 易知,事件的概率具有如下性质:

(1) $P(\overline{A})=1-P(A), P(\varnothing)=0$.

(2) 若$A,B\in\mathscr{F}$,且$B\subset A$,则$P(A-B)=P(A)-P(B)$,且$P(A)\geqslant P(B)$.

(3) 若$A_i\in\mathscr{F}(i=1,2,\cdots,n), A_i\cap A_j=\varnothing(i\neq j)$,则

$$P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)$$

(4) 上连续性:若$A_i\in\mathscr{F}$,且$A_1\supset A_2\supset\cdots$,且$\bigcap_{n=1}^{+\infty}A_n=A$,则

$$\lim_{n\to+\infty}P(A_n)=P(A)$$

(5) 下连续性:若$A_i\in\mathscr{F}$,且$A_1\subset A_2\subset\cdots$,且$\bigcup_{n=1}^{+\infty}A_n=A$,则

$$\lim_{n\to+\infty}P(A_n)=P(A)$$

(6) 设$A_i\in\mathscr{F}, i=1,2,\cdots,n$,有

$$P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i\leqslant k\leqslant n}P(A_iA_k)+\cdots+(-1)^{n+1}P(\bigcap_{i=1}^{n}A_i)$$

(7) 次可加性:若$A_i\in\mathscr{F}, i=1,2,\cdots,n$,则

$$P(\bigcup_{i=1}^{n}A_i)\leqslant\sum_{i=1}^{n}P(A_i)$$

1.1.1 条件概率

定义 1.3 设$(\Omega,\mathscr{F},P)$是概率空间,$A,B\in\mathscr{F}$,且$P(B)>0$,则

$$P(A|B)\triangleq\frac{P(AB)}{P(B)}$$

称为在已知事件B发生的条件下,事件A发生的条件概率.

显然,给定事件B,条件概率$P(\cdot|B)$具有如下性质:

(1) $\forall A \in \mathscr{F}, 0 \leqslant P(A|B) \leqslant 1$；

(2) $P(\Omega|B) = 1$；

(3) 若 $A_i \in \mathscr{F}(i = 1,2,\cdots)$，且 $A_i \cap A_j = \varnothing(i \neq j)$，则

$$P(\bigcup_{i=1}^{+\infty} A_i|B) = \sum_{i=1}^{+\infty} P(A_i|B)$$

定理 1.1 设$(\Omega,\mathscr{F},P)$是概率空间，若 $A_i \in \mathscr{F}$, $P(A_i) > 0$ $(i = 1,2,\cdots)$，$\bigcup_{i=1}^{+\infty} A_i = \Omega, A_iA_j = \varnothing(i \neq j)$. 则对任意 $B \in \mathscr{F}$ 有

(1) $P(B) = \sum_{i=1}^{+\infty} P(A_i)P(B|A_i)$；

(2) $P(A_j|B) = \dfrac{P(A_j)P(B|A_j)}{\sum_{i=1}^{+\infty} P(A_i)P(B|A_i)} (j = 1,2,\cdots)$.

等式(1) 称为全概率公式，等式(2) 称为 Bayes 公式，$P(A_i)$ 称为先验概率，条件概率 $P(A_i \mid B)$ 称为后验概率.

1.1.2 事件的独立性

定义 1.4 设$(\Omega,\mathscr{F},P)$是概率空间，$A,B \in \mathscr{F}$，称 A 与 B 独立，如果

$$P(A \cap B) = P(A)P(B)$$

由定义可知，若事件 A 与 B 独立，则 $\overline{A}$ 与 B，A 与 $\overline{B}$，$\overline{A}$ 与 $\overline{B}$ 也是独立的.

定义 1.5 设$(\Omega,\mathscr{F},P)$是概率空间，$A_1,A_2,\cdots,A_n \in \mathscr{F}$，称 $A_1,A_2,\cdots,A_n$ 是相互独立的，如果对任意$\{A_{i_1},A_{i_2},\cdots,A_{i_m}\} \subset \{A_1,A_2,\cdots,A_n\}(m \leqslant n)$，有

$$P(A_{i_1},A_{i_2},\cdots,A_{i_m}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_m})$$

【例 1.1】 某个国家总人口中有 1% 的人 HIV 呈阳性. 现通过一种程序来检测某人 HIV 是否为阳性，如果被检测的人 HIV 是阳性，通过该程序检测也是阳性的概率是 0.98；如果被检测的人 HIV 不是阳性，通过该程序检测也不是阳性的概率是 0.96. 问：通过该程序检测 HIV 呈阳性，而被检测者 HIV 也是呈阳性的概率有多大？

解 A 表示“该程序检测 HIV 呈阳性”，B 表示“被检测者 HIV 呈阳性”，则由题意可知

$$P(B) = 0.01, P(\overline{B}) = 0.99$$

$$P(A \mid B) = 0.98, P(\overline{A} \mid B) = 0.02$$

$$P(\overline{A} \mid \overline{B}) = 0.96, P(A \mid \overline{B}) = 0.04$$

根据全概率公式，可得

$$P(A) = P(A \mid B)P(B) + P(A \mid \bar{B})P(\bar{B})$$
$$= 0.98 \times 0.01 + 0.04 \times 0.99 = 0.0494$$

再由 Bayes 公式，可知

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)} = \frac{0.98 \times 0.01}{0.0494} = 0.1984$$

1.2 随机变量及其分布

定义 1.6 设$(\Omega,\mathscr{F},P)$是概率空间，$X(\omega)$是定义在Ω上的单值实函数，若对于任意实数$x \in \mathbf{R}$，有$\{\omega: X(\omega) \leqslant x\} \in \mathscr{F}$，则称$X(\omega)$是$\mathscr{F}$上的**随机变量**，简称为**随机变量**. 函数$F(x) = P\{\omega: X(\omega) \leqslant x\}$，$x \in \mathbf{R}$，称为随机变量$X$的**分布函数**.

若随机变量X的可能取值仅有有限个或可列无穷多个，则称X是**离散型随机变量**. 其概率分布用分布列描述：

$$p_k = P\{X = x_k\}, k = 1,2,\cdots$$

且满足$p_k \geqslant 0$，$\sum_{k=1}^{\infty} p_k = 1$，其分布函数$F(x) = \sum_{x_k \leqslant x} p_k$.

对随机变量X的分布函数$F(x)$，若存在一非负函数$f(x)$，对$\forall x \in \mathbf{R}$，有

$$F(x) = \int_{-\infty}^{x} f(u)\mathrm{d}u$$

则称X是**连续型随机变量**，$f(x)$为随机变量X的概率密度函数.

若$f(x)$连续，则

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x)$$

即

$$\lim_{h \to 0} \frac{P(x < X \leqslant x+h)}{h} = f(x)$$

或

$$P(x < X \leqslant x+h) = f(x)h + o(h)$$

以上关系是以后用所谓"微元法"求概率密度函数的依据：为求随机变量X的概率密度函数，先求X落在一个小区域$(x, x+h]$上的概率$P(x < X \leqslant x+h)$，然后令$h \to 0$，求其极限

$$\lim_{h \to 0} \frac{P(x < X \leqslant x+h)}{h}$$

即得$f(x)$.

分布函数具有如下性质：

(1)$F(x)$ 是单调不减函数；

(2)$0 \leqslant F(x) \leqslant 1, \lim_{x \to -\infty} F(x) = 0, \lim_{x \to +\infty} F(x) = 1$；

(3)$F(x)$ 是右连续函数，即对 $\forall x \in \mathbf{R}, F(x+0) = F(x)$.

下面是常见随机变量的分布：

1. 二点分布(伯努利分布)

如果随机变量 X 的分布列为

$$P(X = k) = p^k q^{1-k}, k = 0,1$$

即

$$P(X = 1) = p, P(X = 0) = q = 1 - p$$

其中 $0 < p < 1$,则称 X 服从二点分布或伯努利分布.

2. 几何分布

若随机变量 X 的分布列为

$$P(X = k) = p(1-p)^{k-1} \quad (k = 1,2,\cdots)$$

其中 $0 < p < 1$,则称 X 服从参数为 p 的几何分布,记为 $X \sim g(k;p)$.

3. 泊松分布

如果随机变量 X 的分布列为

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (k = 0,1,2,\cdots)$$

其中 $\lambda > 0$,则称 X 服从参数为 λ 的泊松分布,记为 $X \sim P(\lambda)$.

4. 二项分布

若随机变量 X 的分布列为

$$P(X = k) = C_n^k p^k (1-p)^{n-k} \quad (k = 0,1,\cdots,n)$$

其中 $0 < p < 1$,则称 X 服从参数为 n,p 的二项分布,记为 $X \sim B(k;n,p)$ 或 $X \sim B(n,p)$.

5. 巴斯卡分布

若随机变量 X 的分布列为

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad (k = r, r+1,\cdots)$$

其中 $0 < p < 1, r \in \mathbf{Z}^+$,则称 X 服从参数为 p,r 的巴斯卡分布,记为 $X \sim f(k;r,p)$.

6. 负二项分布

若随机变量 X 的分布列为

$$P(X = l) = \binom{r+l-1}{l} p^r (1-p)^l \quad (l = 0,1,\cdots)$$

其中 $0 < p < 1, r \in \mathbf{Z}^+$,则称 X 服从参数为 p,r 的负二项分布,记为 $X \sim \mathrm{Nb}(l;r,p)$.

7. 超几何分布

若随机变量 X 的分布列为

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\quad k=0,1,\cdots,\min(n,M)$$

其中 $M\leqslant N,n\leqslant N$ 且均为正数，则称 X 服从参数为 M,N,n 的超几何分布，记为 $X\sim h(k;M,N,n)$.

8. 均匀分布

若随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a\leqslant x\leqslant b\\ 0, & \text{其他}\end{cases}$$

则称 X 在区间 $[a,b]$ 上服从均匀分布.

9. 正态分布

若随机变量 X 的概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\quad(-\infty<x<+\infty)$$

其中 μ 和 $\sigma>0$ 是常数，则称随机变量 X 服从参数为 μ 和 σ 的正态分布(或高斯分布)，记为 $X\sim N(\mu,\sigma^2)$. 称参数 $\mu=0,\sigma^2=1$ 的正态分布为标准正态分布，记为 $X\sim N(0,1)$.

10. 对数正态分布

若随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma x}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

其中 $\sigma>0,\mu$ 为常数，则称 X 服从对数正态分布. 因 X 的对数 $\ln X\sim N(\mu,\sigma^2)$ 而得名.

11. 韦布尔分布

若随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{m}{\beta}(x-\alpha)^{m-1}e^{-\frac{(x-\alpha)^m}{\beta}}, & x\geqslant\alpha\\ 0, & x<\alpha\end{cases}$$

则称 X 服从韦布尔分布. 其中 $m(m>0)$ 称为形状参数；α 称为位置参数；$\beta(\beta>0)$ 称为尺度参数.

12. Γ 分布

若随机变量 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}$$

其中 $\alpha>0$，$\lambda>0$，则称 X 服从 Γ 分布，记为 $X\sim\Gamma(\alpha,\lambda)$.

13. **埃尔兰分布**

在 Γ 分布中令 $\alpha=n,\lambda>0$，得到埃尔兰分布的密度函数为

$$f(x)=\begin{cases}\dfrac{\lambda^{n}}{n!}x^{n-1}\mathrm{e}^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}$$

分布函数为

$$F(x)=1-\mathrm{e}^{-\lambda x}\sum_{k=0}^{n-1}\frac{(\lambda x)^k}{k!}=\mathrm{e}^{-\lambda x}\sum_{k=n}^{\infty}\frac{(\lambda x)^k}{k!},x\geqslant 0$$

14. **指数分布**

在 Γ 分布中令 $\alpha=1,\lambda>0$，得到指数分布的密度函数为

$$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}$$

15. χ^2 **分布**

在 Γ 分布中令 $\alpha=\dfrac{n}{2}$，n 为正整数，$\lambda=\dfrac{1}{2}$，得到自由度为 n 的 χ^2 分布的密度函数为

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}x^{\frac{n}{2}}\mathrm{e}^{-\frac{x}{2}}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

16. Beta **分布**

若随机变量 X 的概率密度函数为

$$f(x)=\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1},0\leqslant x\leqslant 1$$

其中参数 $\alpha>0,\beta>0$，则称随机变量 X 服从参数为 α 和 β 的 Beta 分布.

17. Weibull **分布**

若随机变量 X 的概率密度函数为

$$f(x)=\frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1}\mathrm{e}^{(x/\theta)^{\beta}},x>0$$

其中 $\theta>0,\beta>0$，则称随机变量 X 服从参数为 θ 和 β 的 Weibull 分布.

18. Cauch **分布**

若随机变量 X 的概率密度函数为

$$f(x)=\frac{\lambda}{\pi[\lambda^2+(x-\mu)^2]}, -\infty<x<+\infty$$

其中$\lambda>0$，$-\infty<\mu<+\infty$，则称随机变量X服从参数为λ和μ的Cauch分布.

19. Logistic 分布

若随机变量X的概率密度函数为

$$f(x)=\frac{\pi\exp\left(-\frac{\pi}{\sqrt{3}}\cdot\frac{x-\mu}{\sigma}\right)}{\sqrt{3}\sigma\left[1+\exp\left(-\frac{\pi}{\sqrt{3}}\cdot\frac{x-\mu}{\sigma}\right)\right]^2}, -\infty<x<+\infty$$

其中$\sigma>0$，$-\infty<\mu<+\infty$，则称随机变量X服从参数为σ和μ的Logistic分布.

在二项分布中，当n大，p小，而乘积np大小适中时，有如下近似：

定理 1.2　对于二项分布$B(k;n,p)$，若$\lim\limits_{n\to+\infty}np=\lambda$，则

$$\lim_{n\to+\infty}C_n^k p^k q^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}$$

由此可得，当n充分大，p较小时，二项分布可用泊松分布近似表示.

定理 1.3　当N很大，而n很小时，则有

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\approx\binom{n}{m}p^m(1-p)^{n-m}, \ p=\frac{M}{N}$$

从而当N很大，而n很小时，可用二项分布来近似表示超几何分布.

【例 1.2】　一般来说，只有0.01%的鳟鱼卵可以发育成鳟鱼，那么40000个鳟鱼卵中至少有3个可以发育成鳟鱼的概率是多少？

解　令X表示40000个鳟鱼卵中发育成鳟鱼的个数，则X服从$p=0.01\%$的二项分布，$p_i=P(X=i)=\binom{40000}{i}p^i(1-p)^{40000-i}$，$i=1,2,\cdots,40000$.

又n较大，p较小，故可用$\lambda=np=4$的泊松分布近似表示二项分布，即

$$p_i\approx\frac{4^i}{i!}e^{-4}, i=0,1,\cdots,40000$$

所以40000个鳟鱼卵中至少有3个可以发育成鳟鱼的概率为

$$\begin{aligned}p=P(X\geqslant 3)&=1-p_0-p_1-p_2\\&\approx 1-0.0183-0.0733-0.1465=0.7619\end{aligned}$$

定义 1.7 $(\Omega,\mathscr{F},P)$ 是概率空间，$\boldsymbol{X}=\boldsymbol{X}(\omega)=(X_1(\omega),X_2(\omega),\cdots,X_n(\omega))$ 是定义在 Ω 上取值于 n 维空间 $\mathbf{R}^n$ 的向量函数. 如果 $X_i(\omega)$，$1\leqslant i\leqslant n$ 是 $(\Omega,\mathscr{F},P)$ 上的随机变量，则称 $\boldsymbol{X}=\boldsymbol{X}(\omega)$ 为 n 维随机变量或 n 维随机向量，且称

$$F(x)=F(x_1,x_2,\cdots,x_n)=P\{\omega:X_1(\omega)\leqslant x_1,X_2(\omega)\leqslant x_2,\cdots,X_n(\omega)\leqslant x_n\},$$

$$x=(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n$$

为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的联合分布函数.

特别地，$F(x,y)=P\{X\leqslant x,Y\leqslant y\}$，$x,y\in\mathbf{R}$ 是 X,Y 的二维联合分布函数.

如果 X,Y 的可能取值仅有有限对或可列无穷多对，则称 (X,Y) 是二维离散型随机向量. 设 X,Y 所有可能取值为 (x_i,y_j)，$i=1,2,\cdots,j=1,2,\cdots$，则 (X,Y) 的联合概率分布可用下面的联合分布律表示：

$$p_{ij}=P\{X=x_i,Y=y_j\},\ i=1,2,\cdots,j=1,2,\cdots$$

且满足 $p_{ij}\geqslant 0$，$\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty}p_{ij}=1$.

如果存在一个非负的二元函数 $f(x,y)$，使得任意 $x,y\in\mathbf{R}$，有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(x,y)\mathrm{d}x\mathrm{d}y$$

则称 (X,Y) 是二维连续型随机向量，$f(x,y)$ 称为 (X,Y) 的联合密度函数，它具有下列性质：

(1) $f(x,y)\geqslant 0$；

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1$；

(3) 在 $f(x,y)$ 的连续点 (x,y) 处，有

$$f(x,y)=\frac{\partial^2F(x,y)}{\partial x\partial y}$$

由定义 1.7 知，$X_i(\omega)$ 的分布函数可由联合分布函数求得，即 $F_{X_i}(x_i)=F(\infty,\cdots,\infty,x_i,\infty,\cdots,\infty)$，称一维分布函数 $F_{X_1}(x_1),F_{X_2}(x_2),\cdots,F_{X_n}(x_n)$ 为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的边缘分布函数. 反之，随机向量的边缘分布函数不能确定其联合分布函数，但当它们是相互独立时，边缘分布和联合分布包含着相同的信息.

定义 1.8 $(X_1,X_2,\cdots,X_n)$ 是 n 维随机变量，若 $\forall(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n$，有

$$F(x_1,x_2,\cdots,x_n)=F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$$

则称随机变量 $X_1,X_2,\cdots,X_n$ 相互独立.

定理 1.4 若随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，则其中任意 $k(2\leqslant k\leqslant n)$ 个随机变量 $X_{i_1},X_{i_2},\cdots,X_{i_k}$ 也相互独立.

1.3 随机变量的函数及其分布

先讨论单个随机变量函数的分布,设X是连续型随机变量,$y=h(x)$是一实值函数.

定理1.5 若$Y=h(X)=\alpha X+\beta$,则

当$\alpha>0$时,$F_Y(y)=F_X(\frac{y-\beta}{\alpha})$;

当$\alpha<0$时,$F_Y(y)=1-F_X(\frac{y-\beta}{\alpha})$;

当$\alpha\neq 0$时,$f_Y(y)=\left|\frac{1}{\alpha}\right|f_X(\frac{y-\beta}{\alpha})$.

定理1.6 若$y=h(x)$是一个严格单调的函数,逆函数为$x=h^{-1}(y)$,则

(1) 当$y=h(x)$是严格单调递增函数时,$F_Y(y)=F_X[h^{-1}(y)]$;

(2) 当$y=h(x)$是严格单调递减函数时,$F_Y(y)=1-F_X[h^{-1}(y)]$;

(3) $f_Y(y)=f_X[h^{-1}(y)]\left|\frac{dh^{-1}(y)}{dy}\right|=f_X[x(y)]\left|\frac{dx(y)}{dy}\right|$.

【例1.3】 一个质量为m的质点以随机速度X沿着直线移动,已知$X\sim U[0,V]$,该质点随机运动能量$Y=\frac{1}{2}mX^2$,求Y的分布.

解 由$y=h(x)=\frac{1}{2}mx^2$,$x\geqslant 0$知,

$x=h^{-1}(y)=\sqrt{\frac{2y}{m}}$, $\frac{dx}{dy}=\sqrt{\frac{1}{2my}}$,$0<y<\frac{1}{2}mV^2$.

而$f_X(x)=\frac{1}{V}$,$0\leqslant x\leqslant V$,所以$f_Y(y)=\frac{1}{V}\sqrt{\frac{1}{2my}}$,$0<y<\frac{1}{2}mV^2$.

【例1.4】 设随机变量$X\sim N(0,1)$,求$Y=e^{3X+1}$的概率密度函数.

解 $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$,$-\infty<x<+\infty$

因$Y=e^{3X+1}$,所以函数$y=e^{3x+1}$单调递增.而

$$x=\frac{1}{3}\ln y-\frac{1}{3},\quad x'=\frac{1}{3y}$$

故Y的概率密度为

$$f_Y(y)=\begin{cases}\varphi\left(\frac{1}{3}\ln y-\frac{1}{3}\right)\cdot x'=\frac{1}{3y\sqrt{2\pi}}e^{-\frac{(\frac{1}{3}\ln y-\frac{1}{3})^2}{2}}, & y\geqslant 0\\ 0, & y<0\end{cases}$$

1.3.1 二维随机变量函数的分布

已知二维随机变量(X,Y)的联合分布，如何求$g(X,Y)$的分布？下面就连续型随机变量进行讨论.

设连续型随机变量(X,Y)的联合概率密度为$f(x,y)$，则随机变量$Z=g(X,Y)$的分布函数为

$$F_Z(z)=\iint\limits_{g(x,y)\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y \quad (-\infty<z<+\infty)$$

Z的概率密度为

$$f_Z(z)=\frac{\mathrm{d}F_Z(z)}{\mathrm{d}z} \quad (-\infty<z<+\infty)$$

1. 和$Z=X+Y$的分布

已知(X,Y)的联合概率密度为$f(x,y)$，则$Z=X+Y$的分布函数为

$$\begin{aligned}F_Z(z)&=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}\\&=\iint\limits_{x+y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-x}f(x,y)\mathrm{d}y\right]\mathrm{d}x\end{aligned}$$

令$y=u-x$，得

$$F_Z(z)=\int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty}f(x,u-x)\mathrm{d}x\right]\mathrm{d}u$$

于是，Z的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x$$

利用X与Y的对称性可得

$$f_Z(z)=\int_{-\infty}^{+\infty}f(z-y,y)\mathrm{d}y$$

当X与Y相互独立时，得卷积公式

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)\mathrm{d}y$$

2. 商$\frac{X}{Y}$的分布

已知(X,Y)的联合概率密度为$f(x,y)$，则$Z=\frac{X}{Y}(Y\neq 0)$的分布函数为

$$\begin{aligned}F_Z(z)&=P\{X/Y\leqslant z\}=\iint\limits_{x/y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y\\&=\int_{0}^{+\infty}\left[\int_{-\infty}^{yz}f(x,y)\mathrm{d}x\right]\mathrm{d}y+\int_{-\infty}^{0}\left[\int_{yz}^{+\infty}f(x,y)\mathrm{d}x\right]\mathrm{d}y\end{aligned}$$

令$x=uy$，得

$$F_Z(z)=\int_{-\infty}^{z}\left[\int_{0}^{+\infty}yf(yu,y)\mathrm{d}y-\int_{-\infty}^{0}yf(yu,y)\mathrm{d}y\right]\mathrm{d}u$$

于是

$$f_Z(z)=\int_{0}^{+\infty}yf(yz,y)\mathrm{d}y-\int_{-\infty}^{0}yf(yz,y)\mathrm{d}y=\int_{-\infty}^{+\infty}|y|f(yz,y)\mathrm{d}y$$

X 与 Y 相互独立时，

$$f_Z(z)=\int_{-\infty}^{+\infty}|y|f_X(yz)f_Y(y)\mathrm{d}y$$

3. $M=\max(X,Y)$ **及** $N=\min(X,Y)$ **的分布**

设 X 与 Y 相互独立，则

$$\begin{aligned}F_{\max}(z)&=P\{M\leqslant z\}=P\{X\leqslant z,Y\leqslant z\}=P\{X\leqslant z\}\cdot P\{Y\leqslant z\}\\&=F_X(z)F_Y(z)\end{aligned}$$

$$\begin{aligned}F_{\min}(z)&=P\{N\leqslant z\}=1-P\{N>z\}=1-P\{X>z,Y>z\}\\&=1-P\{X>z\}\cdot P\{Y>z\}=1-[1-P\{X\leqslant z\}][1-P\{Y\leqslant z\}]\\&=1-[1-F_X(z)][1-F_Y(z)]\end{aligned}$$

推广：设 $X_1,X_2,\cdots,X_n$ 是 n 个相互独立的随机变量，它们的分布函数为 $F_{X_i}(x_i)(i=1,2,\cdots,n)$，则 $M=\max\{X_1,X_2,\cdots,X_n\}$ 及 $N=\min\{X_1,X_2,\cdots,X_n\}$ 的分布函数分别为

$$F_{\max}(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z)$$

$$F_{\min}(z)=1-[1-F_{X_1}(z)][1-F_{X_2}(z)]\cdots[1-F_{X_n}(z)]$$

特别地，当 $X_1,X_2,\cdots,X_n$ 相互独立且具有相同分布函数 $F(x)$ 时有

$$F_{\max}(z)=[F(z)]^n,F_{\min}(z)=1-[1-F(z)]^n$$

【例 1.5】 设 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\mathrm{e}^{-(x+y)}, & x>0,y>0\\0, & \text{其他}\end{cases}$$

试求 $Z=X+Y$ 和 $W=\dfrac{X}{Y}$ 的概率密度.

解 当 $z\leqslant 0$ 时，$z-x\leqslant 0$，$f(x,z-x)\equiv 0$，所以

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x=0$$

当 $z>0$ 时，$x>0$，$z-x>0$，$f(x,z-x)=\mathrm{e}^{-(x+z-x)}$

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x=\int_{0}^{z}\mathrm{e}^{-(x+z-x)}\mathrm{d}x=z\mathrm{e}^{-z}$$

故 $Z=X+Y$ 的概率密度为

$$f_Z(z)=\begin{cases}z\mathrm{e}^{-z}, & z>0\\0, & \text{其他}\end{cases}$$

又当 $z>0$ 时，

$$f_W(z)=\int_{-\infty}^{+\infty}|y|f(zy,y)\mathrm{d}y=\int_0^{+\infty}y\mathrm{e}^{-(zy+y)}\mathrm{d}y=\frac{1}{(z+1)^2}$$

当 $z\leqslant 0$ 时，$f(zy,y)=0$，故 $f_W(z)=0$.

故 $W=\dfrac{X}{Y}$ 的概率密度为

$$f_W(z)=\begin{cases}\dfrac{1}{(z+1)^2}, & z>0\\ 0, & z\leqslant 0\end{cases}$$

1.3.2　随机向量变换的分布

定理 1.7　设 (X_1,X_2) 的联合密度为 $f(x_1,x_2)$，若函数

$$\begin{cases}y_1=g_1(x_1,x_2)\\ y_2=g_2(x_1,x_2)\end{cases}$$

满足下述条件：

① 存在唯一反函数 $\begin{cases}x_1=x_1(y_1,y_2)\\ x_2=x_2(y_1,y_2)\end{cases}$；

② 有连续的一阶偏导数；

③Jacobi 行列式 $J=\begin{vmatrix}\dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2}\\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2}\end{vmatrix}\neq 0$.

则 $Y_1=g_1(X_1,X_2)$，$Y_2=g_2(X_1,X_2)$ 的联合概率密度为

$$f[x_1(y_1,y_2),x_2(y_1,y_2)]|J|$$

【例 1.6】　设 X_1,X_2 的联合密度为 $f(x_1,x_2)$，而 $Y_1=aX_1+bX_2$，$Y_2=cX_1+dX_2$，$A=\begin{vmatrix}a & b\\ c & d\end{vmatrix}\neq 0$，试求 (Y_1,Y_2) 的联合密度函数 $q(y_1,y_2)$.

解　由 $y_1=ax_1+bx_2$，$y_2=cx_1+dx_2$ 知

$$x_1=\frac{d}{A}y_1-\frac{b}{A}y_2,\ x_2=-\frac{c}{A}y_1+\frac{a}{A}y_2$$

而

$$J=\begin{vmatrix}\dfrac{d}{A} & -\dfrac{b}{A}\\ -\dfrac{c}{A} & \dfrac{a}{A}\end{vmatrix}=\frac{ad-bc}{A^2}=\frac{1}{ad-bc}$$

从而

$$q(y_1,y_2)=\frac{f\left(\dfrac{d}{A}y_1-\dfrac{b}{A}y_2,-\dfrac{c}{A}y_1+\dfrac{a}{A}y_2\right)}{|ad-bc|}$$

1.4 矩、数学期望和方差

定义 1.9 ①（m 阶原点矩）离散型随机变量 X 的取值为 $x_1, x_2, \cdots$，如果 $\sum_{i=1}^{+\infty} |x_i|^m P\{X = x_i\} < +\infty$，则称

$$E(X^m) \triangleq \sum_{i=1}^{+\infty} x_i^m P\{X = x_i\}$$

为 X 的 m 阶原点矩. 当 $m = 1$ 时，称为 X 的数学期望或均值.

② 连续型随机变量 X 的分布函数为 $F(x)$，其密度函数为 $f(x)$，若 $\int_{-\infty}^{+\infty} |x|^m f(x)\mathrm{d}x < +\infty$，则称

$$E(X^m) = \int_{-\infty}^{+\infty} x^m F'(x)\mathrm{d}x = \int_{-\infty}^{+\infty} x^m f(x)\mathrm{d}x$$

为 X 的 m 阶矩. X 的一阶矩称为 X 的数学期望或均值，记为 μ_X 或 m_X.

定义 1.10 （m 阶中心矩）若 X 的数学期望 μ_X 存在，则称

$$c_m = E(X - \mu_X)^m$$

为 X 的 m 阶中心矩. X 的二阶中心矩称为 X 的方差，记为 $\mathrm{Var}(X)$ 或 σ_X^2，称 $\sigma_X = \sqrt{D(X)}$ 为 X 的标准差.

定理 1.8 设 X 是连续型随机变量，$y = h(x)$ 是一实值函数，则

$$E[h(X)] = \begin{cases} \sum_{i=1}^{+\infty} h(x_i)P\{X = x_i\}, \text{当 } X \text{ 是离散型随机变量} \\ \int_{-\infty}^{+\infty} h(x)f_X(x)\mathrm{d}x, \text{当 } X \text{ 是连续型随机变量} \end{cases}$$

随机变量的期望和方差的性质：

(1) $E(C) = C, D(C) = 0$；

(2) $E(CX) = CE(X), D(CX) = C^2 D(X)$；

(3) $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$；

(4) 若 X 和 Y 相互独立，则 $E(XY) = E(X)E(Y)$；

(5) $D(X) = 0$ 的充要条件是 $P\{X = E(X)\} = 1$；

(6) 若 X 为非负整数值的随机变量，则

$$E(X) = \sum_{i=1}^{+\infty} P\{X \geqslant i\}$$

(7) 若 X 为连续型随机变量，则

$$E(X) = \int_0^{+\infty} [1 - F(x)]\mathrm{d}x$$

定义 1.11 对于两个随机变量 X,Y，称

$$\mathrm{Cov}(X,Y) = E\{[X - E(X)][Y - E(Y)]\} = E(XY) - E(X)E(Y)$$

为 (X,Y) 的**协方差**.

由上式知，若 X,Y 相互独立，则 $E(XY) = E(X)E(Y)$，从而得 $\mathrm{Cov}(X,Y) = 0$. 于是，若 $\mathrm{Cov}(X,Y) \neq 0$，则 X,Y 不独立. 因此，$\mathrm{Cov}(X,Y) \neq 0$ 刻画了 X,Y 的取值存在某种统计上的线性相关关系.

定义 1.12 X,Y 是随机变量，若 $0 < D(X) = \sigma_X^2 < +\infty$，$0 < D(Y) = \sigma_Y^2 < +\infty$，称

$$r_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

为 (X,Y) 的**相关系数**.

相关系数具有如下性质：

(1) 若 X 和 Y 相互独立，则 $r_{XY} = 0$.

(2) $|r_{XY}| = 1$ 当且仅当存在常数 a 和 b，使得

$$P\{Y = aX + b\} = 1$$

(3) 对任意的随机变量 X,Y，$-1 \leqslant r_{XY} \leqslant 1$.

(4) $D(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} D(X_i) + 2\sum_{\substack{i,j=1\\ i<j}}^{n} \mathrm{Cov}(X_i, X_j)$.

(5)(Schwarz 不等式) 若 $E(X^2) < +\infty$，$E(Y^2) < +\infty$，则

$$E(|XY|)^2 \leqslant E(X^2)E(Y^2)$$

特别地

$$|\mathrm{Cov}(X,Y)|^2 \leqslant \sigma_X^2 \sigma_Y^2$$

(6) $\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j}\sum \mathrm{Cov}(X_i, X_j)$.

r_{XY} 刻画了 X,Y 之间线性关系的密切程度，若 $r_{XY} = 0$，则称 X,Y **不相关**. 由定义可知，X,Y 不相关当且仅当

$$E(XY) = E(X)E(Y)$$

因此，若 X,Y 独立，则它们不相关. 但是如果 X,Y 不相关，则它们不一定独立. 看下面的例子：

【例 1.7】 (匹配问题) 在一次集会上，n 个人把他们的帽子放到房间的中央混

合在一起，然后每个人随机地选取一顶，求拿到自己帽子的人数 X 的均值和方差.

解　利用表达式 $X = X_1 + X_2 + \cdots + X_n$，其中

$$X_i = \begin{cases} 1, & \text{如果第 } i \text{ 个人拿到自己的帽子} \\ 0, & \text{其他} \end{cases}$$

即求 $E(X)$、$D(X)$，因为 $P\{X_i = 1\} = \frac{1}{n}$，所以

$$E(X_i) = \frac{1}{n}, D(X_i) = \frac{1}{n} - (\frac{1}{n})^2 = \frac{n-1}{n^2}$$

又

$$\mathrm{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$$

而

$$X_i X_j = \begin{cases} 1, & \text{如果第 } i \text{ 个人与第 } j \text{ 个人都拿到自己的帽子} \\ 0, & \text{其他} \end{cases}$$

于是

$$\begin{aligned} E(X_i X_j) &= P\{X_i = 1, X_j = 1\} \\ &= P\{X_i = 1\} P\{X_j = 1 \mid X_i = 1\} = \frac{1}{n} \cdot \frac{1}{n-1} \end{aligned}$$

故

$$\mathrm{Cov}(X_i, X_j) = \frac{1}{n(n-1)} - \left(\frac{1}{n}\right)^2 = \frac{1}{n^2(n-1)}$$

所以

$$E(X) = 1, \quad D(X) = \frac{n-1}{n} + 2\mathrm{C}_n^2 \frac{1}{n^2(n-1)} = 1$$

本题是将 X 分解成数个随机变量之和，然后利用随机变量和的数学期望等于随机变量数学期望之和来求 $E(X)$，这种处理方法在实际中常常用到.

【例 1.8】　随机变量 X 的取值为 $-2, -1, 1, 2$，随机变量 Y 的取值为 $-1, 0, 1$. 它们的联合分布律见表 1.1. 试证 X, Y 不相关，但也不独立.

表 1.1　X, Y 的联合分布律

X \ Y	-1	0	1
-2	1/16	1/16	1/16
-1	2/16	1/16	2/16
1	2/16	1/16	2/16
2	1/16	1/16	1/16

证明

$$E(Y)=\frac{6}{16}\cdot(-1)+\frac{4}{16}\cdot 0+\frac{6}{16}\cdot 1=0$$

$$E(X)=\frac{3}{16}\cdot(-2)+\frac{5}{16}\cdot(-1)+\frac{5}{16}\cdot 1+\frac{3}{16}\cdot 2=0$$

$$\begin{aligned}E(XY)=&\frac{1}{16}\cdot(-2)\cdot(-1)+\frac{1}{8}\cdot(-1)\cdot(-1)+\frac{1}{8}\cdot 1\cdot(-1)+\frac{1}{16}\cdot 2\cdot(-1)\\&+\frac{1}{16}\cdot(-2)\cdot 0+\frac{1}{16}\cdot(-1)\cdot 0+\frac{1}{16}\cdot 1\cdot 0+\frac{1}{16}\cdot 2\cdot 0\\&+\frac{1}{16}\cdot(-2)\cdot 1+\frac{1}{8}\cdot(-1)\cdot 1+\frac{1}{8}\cdot 1\cdot 1+\frac{1}{16}\cdot 2\cdot 1=0\end{aligned}$$

所以 $E(XY)=E(X)E(Y)$，从而 X,Y 不相关.

另一方面，

$$P\{X=2,Y=-1\}=\frac{1}{16}\neq P\{X=2\}P\{Y=-1\}=\frac{3}{16}\cdot\frac{6}{16}=\frac{18}{256}=\frac{9}{128}$$

因此 X,Y 不相关，但不独立.

【例 1.9】 随机向量 (X,Y) 的联合概率密度函数为

$$f_{X,Y}(x,y)=\frac{x^2+y^2}{4\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\},\quad -\infty<x<y<+\infty$$

试证明 X,Y 不相关，但不独立.

证明

$$\begin{aligned}f_X(x)&=\int_{-\infty}^{+\infty}f_{X,Y}(x,y)\mathrm{d}y=\int_{-\infty}^{+\infty}\frac{x^2+y^2}{4\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\}\mathrm{d}y\\&=\frac{\mathrm{e}^{-x^2/2}}{2\sqrt{2\pi}}\left(x^2\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2}\mathrm{d}y+\int_{-\infty}^{+\infty}y^2\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2}\mathrm{d}y\right)\\&=\frac{1}{2\sqrt{2\pi}}(x^2+1)\mathrm{e}^{-x^2/2},-\infty<x<+\infty\end{aligned}$$

第二行等式括号里的第一个被积函数为标准正态分布的密度函数，第二个积分为标准正态分布的方差.

因为 $f_{X,Y}(x,y)$ 关于 x,y 是对称的，故

$$f_Y(y)=\frac{1}{2\sqrt{2\pi}}(y^2+1)\mathrm{e}^{-y^2/2},-\infty<y<+\infty$$

很明显有 $f_{X,Y}(x,y)\neq f_X(x)f_Y(y)$，因此 X,Y 不独立.

$f_X(x)$ 和 $f_Y(y)$ 都是关于原点对称的，从而有 $E(X)=E(Y)=0$，

$$\begin{aligned}E(XY)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\frac{x^2+y^2}{4\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\}\mathrm{d}x\mathrm{d}y\\&=\frac{1}{4\pi}\left(\int_{-\infty}^{+\infty}x^3\mathrm{e}^{-x^2/2}\mathrm{d}x\right)\left(\int_{-\infty}^{+\infty}y^3\mathrm{e}^{-y^2/2}\mathrm{d}x\right)=0\end{aligned}$$

故 $E(XY)=E(X)E(Y)=0$.

到此已经证明了 X,Y 是不相关的，但它们不独立.

【例 1.10】 $X\sim U[0,\pi]$，$Y=\sin X$，证明 X,Y 不相关.

证明 $E(X)=\frac{\pi}{2}$，$E(Y)=\frac{1}{\pi}\int_0^{\pi}\sin x\mathrm{d}x=\frac{2}{\pi}$，$E(XY)=\frac{1}{\pi}\int_0^{\pi}x\sin x\mathrm{d}x=1$.

X,Y 的协方差
$$\mathrm{Cov}(X,Y)=1-\frac{\pi}{2}\cdot\frac{2}{\pi}=0$$

从而 X,Y 是不相关的.

在例 1.10 中 Y 是 X 的函数，它们之间有很强的统计相关关系，但根据相关系数，它们之间是不相关的. 这进一步说明了相关系数只能刻画随机变量间的线性相关性.

【例 1.11】 (X,Y) 是二元正态随机向量，其概率密度函数为

$$f_{X,Y}(x,y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2}-2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}+\frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\}$$

其中参数 $-\infty<\mu_x,\mu_y<+\infty,\sigma_x>0,\sigma_y>0,-1<\rho<1$.

试证：X,Y 独立当且仅当 $\rho=0$.

证明

$$f_X(x)=\int_{-\infty}^{+\infty}f_{X,Y}(x,y)\mathrm{d}y=\frac{1}{\sqrt{2\pi}\sigma_x}\exp\left[-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right],\quad -\infty<x<+\infty$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f_{X,Y}(x,y)\mathrm{d}x=\frac{1}{\sqrt{2\pi}\sigma_y}\exp\left[-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right],\quad -\infty<y<+\infty$$

因为 X,Y 独立等价于 $f_{X,Y}(x,y)=f_X(x)f_Y(y)$，因此 X,Y 独立当且仅当 $\rho=0$.

可以计算 ρ 等于 X,Y 的相关系数，由此可知若 (X,Y) 为二元正态随机向量，则 X,Y 独立当且仅当它们不相关，并且 X,Y 都是正态随机变量.

1.5 条件期望

1.5.1 离散型情形

定义 1.13 (X,Y) 是离散型随机向量，如果 $P\{Y=y\}>0$，给定 $Y=y$ 条件下 X 的条件概率的定义为

$$P_{XY}\{x\mid y\}=P\{X=x|Y=y\}=\frac{P\{X=x,Y=y\}}{P\{Y=y\}}$$

给定 $Y=y$ 条件下 X 的条件分布函数的定义为

$$F_{X|Y}(x|y)=P\{X\leqslant x|Y=y\}=\frac{\sum_{x_i\leqslant x}P\{X=x_i,Y=y\}}{P\{Y=y\}}$$

给定 $Y=y$ 条件下 X 的条件期望的定义为

$$E(X|y)=\sum_i x_i\frac{P\{X=x_i,Y=y\}}{P\{Y=y\}}$$

由定理 1.1 的全概率公式知

$$P\{X=x\}=\sum_y P\{X=x\mid Y=y\}P\{Y=y\}=\sum_y P_{XY}\{x\mid y\}P_Y\{y\}$$

【例 1.12】 X 服从参数为 N,p 的二项分布，其中随机变量 N 是服从参数为 λ 的 Poisson 分布，求 $P\{X=k\}$.

解 由全概率公式可知

$$\begin{aligned}
P\{X=k\}&=\sum_{n=0}^{+\infty}P_{XN}\{x\mid n\}P_N\{n\}\\
&=\sum_{n=k}^{+\infty}\binom{n}{k}p^k(1-p)^{n-k}\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}\\
&=\frac{(\lambda p)^k}{k!}\mathrm{e}^{-\lambda}\sum_{n=k}^{+\infty}\frac{[\lambda(1-p)]^{n-k}}{(n-k)!}\\
&=\frac{(\lambda p)^k}{k!}\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda(1-p)}\\
&=\frac{(\lambda p)^k}{k!}\mathrm{e}^{-\lambda p}
\end{aligned}$$

【例 1.13】 X 服从参数为 N,p 的负二项分布，其中随机变量 N 是服从参数为 β 的几何分布，求 $P\{X=k\}$.

解 由全概率公式可知

$$\begin{aligned}
P\{X=k\}&=\sum_{n=0}^{+\infty}P_{XN}\{x\mid n\}P_N\{n\}\\
&=\sum_{n=k}^{+\infty}\binom{n+k-1}{k}p^n(1-p)^k(1-\beta)\beta^{n-1}\\
&=(1-p)^k(1-\beta)p\sum_{n=1}^{+\infty}\binom{n+k-1}{k}(p\beta)^{n-1}\\
&=(1-p)^k(1-\beta)p(1-p\beta)^{-k-1}\\
&=\frac{p-p\beta}{1-p\beta}\left(\frac{1-p}{1-p\beta}\right)^k
\end{aligned}$$

1.5.2　连续型情形

定义 1.14　连续型随机向量(X, Y)的联合概率密度函数为$f_{XY}(x,y)$，给定$Y=y$条件下，X的条件密度函数为

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$$

给定$Y=y$条件下，X的条件分布函数为

$$F_{X|Y}(x|y)=P\{X\leqslant x|Y=y\}=\int_{-\infty}^{x}f_{X|Y}(u|y)\mathrm{d}u=\frac{\int_{-\infty}^{x}f(u,y)\mathrm{d}u}{f_Y(y)}$$

给定$Y=y$条件下，X的条件数学期望为

$$E(X|y)=E(X|Y=y)\triangleq\int_{-\infty}^{+\infty}x\mathrm{d}F_{X|Y}(x|y)$$

显然，条件期望$E(X|y)$是关于y的函数，y是Y的一个取值. 若在已知Y的条件下，全面考虑X的均值，需要以Y代替y，则$E(X|Y)$既是随机变量Y的函数，也是随机变量，称为X在Y下的条件期望. 且$E(X|Y)$在$Y=y$时，取值为$E(X|Y=y)$.

条件期望在概率论、数理统计和随机过程中是一个十分重要的概念，下面介绍一个极其有用的性质.

定理 1.9　（全期望公式）随机变量X和Y的期望存在，则

$$E(X)=E[E(X|Y)]=\int E(X|Y=y)\mathrm{d}F_Y(y)$$

如果Y是离散型随机变量，则上式为

$$E(X)=\sum_y E(X|Y=y)P\{Y=y\}$$

如果Y是连续型随机变量，概率密度函数为$f(y)$，则上式为

$$E(X)=\int_{-\infty}^{+\infty}E(X|Y=y)f(y)\mathrm{d}y$$

由全期望公式可知，$E(X)$是在给定$Y=y$时，X的条件期望的一个加权平均值，每一项$E(X|Y=y)$所加的权是作为条件的事件的概率.

定理 1.10　设(X, Y)是二维随机变量，条件分布函数$F_{X|Y}(x|y)$存在，函数$g(x)$在$\mathbf{R}$上连续，若$\int_{-\infty}^{+\infty}|g(x)|\mathrm{d}F_{X|Y}(x|y)<+\infty$，则随机变量$g(X)$在“$Y=y$”条件下的条件数学期望为

$$E[g(X)|Y=y]=\int_{-\infty}^{+\infty}g(x)\mathrm{d}F_{X|Y}(x|y)$$

条件期望$E(X \mid Y=y)$是关于条件概率的期望,而条件概率具有概率的性质,因此,条件期望也有很多类似期望的性质。

设X,Y,Z是$(\Omega,\mathscr{F},P)$上的随机变量,$g(\cdot),g_1(\cdot),g_2(\cdot)$和$h(\cdot)$为$\mathbf{R}$上的函数,$v(\cdot)$是二元函数,$c_1,c_2$是常数,且各数学期望均存在,则有

(1)$E\{[c_1g_1(X)+c_2g_2(X)] \mid Y=y\}=c_1E[g_1(X) \mid Y=y]+c_2E[g_2(X) \mid Y=y]$.

(2)$E[c_1 \mid Y=y]=c_1$.

(3)$E[v(X,Y) \mid Y=y]=E[v(X,y) \mid Y=y]$.

(4) 如果X与Y相互独立,则$E[g(X) \mid Y=y]=E[g(X)]$.

(5)$E[g(X)h(Y) \mid X=x]=g(x)E[h(Y) \mid X=x]$.

(6)$E[g(X)h(Y)]=E\{g(X)E[h(Y) \mid X]\}$.

由性质(1),(5),(6)可知,$E[g(X) \mid X=x]=g(x)$;$E[g(X)]=E\{E[g(X) \mid Y]\}$.

【例 1.14】 设随机变量X的数学期望$E(X)$存在,$X=\sum_{k=1}^{N}\eta_k$,其中$\eta_k,k=1,2,\cdots,N$都是随机变量,并且相互独立;N仅取自然数,$E(N)$存在,证明:

$$E(X)=\sum_{k=1}^{+\infty}P\{N \geqslant k\}E(\eta_k)$$

证明

$$\begin{aligned}E(X)=E[E(X \mid N=n)]&=\sum_{n=1}^{+\infty}P\{N=n\}E\Big(\sum_{k=1}^{N}\eta_k \,\Big|\, N=n\Big)\\&=\sum_{n=1}^{+\infty}P\{N=n\}\sum_{k=1}^{n}E(\eta_k)\\&=\sum_{k=1}^{+\infty}E(\eta_k)\sum_{n=k}^{+\infty}P\{N=n\}\\&=\sum_{k=1}^{+\infty}E(\eta_k)P\{N \geqslant n\}\end{aligned}$$

又若η_k具有相同分布,则

$$E(X)=E(\eta_1)\sum_{k=1}^{+\infty}P\{N \geqslant k\}=E(\eta_1)\sum_{k=1}^{+\infty}kP\{N=k\}=E(\eta_1)E(N)$$

先对一个适当的随机变量取条件,不仅可以求得期望,也可以用这种方法计算事件的概率. 设M是一个任意事件,定义示性随机变量为

$$I_M(\omega)=\begin{cases}1, & 若\ \omega \in M\\0, & 若\ \omega \notin M\end{cases}$$

显然有

$$P\{M\} = E[I_M(\omega)] = E\{E[I_M(\omega) \mid Y = y]\}$$
$$= \int E[I_M(\omega) \mid Y = y] \mathrm{d}F_Y(y)$$
$$= \int P\{I_M(\omega) \mid Y = y\} \mathrm{d}F_Y(y)$$

Y 为任意的随机变量.

【例 1.15】 $X \sim F(x)$, $Y \sim G(y)$ 是相互独立的随机变量, $X+Y$ 的分布记为 $F * G$, 称为 F 与 G 的卷积, 则

$$F * G(z) = P\{X+Y \leqslant z\}$$
$$= \int_{-\infty}^{+\infty} P\{X+Y \leqslant z \mid Y = y\} \mathrm{d}G(y)$$
$$= \int_{-\infty}^{+\infty} P\{X \leqslant z-y\} \mathrm{d}G(y)$$

$F * F$ 记为 F_2, $F * F_2$ 记为 F_3, 以此类推, 可得到 $F_n = F * F_{n-1}$, 称为 F 的 n 重卷积, 是 n 个独立且具有共同分布 F 的随机变量的和的分布.

【例 1.16】 (X, Y) 是二维正态分布, 即 $(X,Y) \sim N(\mu_1, \mu_2, \rho, \sigma_1^2, \sigma_2^2)$, 则联合概率密度函数为

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2 (1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

则

$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)}$$
$$= \frac{1}{\sqrt{2\pi}\sigma_2 (1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma_2^2(1-\rho^2)}\left[y-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right]^2\right\}$$

为正态分布的密度函数, 故

$$E(Y \mid X = x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$$

$$E(Y \mid X) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X-\mu_1)$$

【例 1.17】 (匹配问题续) 设有 n 个人, 把他们的帽子混在一起后, 每人随机地选一顶, 求恰好有 k 个人选到自己帽子的概率.

解 记 E 表示全部匹配这一事件及 $P_n = P\{E\}$, 表明依赖于 n, M 为第一个人选到自己帽子这一事件, $\overline{M}$ 为第一个人没有选到自己帽子这一事件. 利用全概率公式得

$$P_n = P\{E\} = P\{E \mid M\}P\{M\} + P\{E \mid \overline{M}\}P\{\overline{M}\}$$

由于 $P\{E \mid M\}P\{M\} = 0$, 从而

$$P_n = P\{E \mid \overline{M}\}P\{\overline{M}\} = P\{E \mid \overline{M}\}\frac{n-1}{n}$$

$P(E \mid \overline{M})$ 是 $n-1$ 个人从 $n-1$ 顶帽子中各取一顶都不匹配的概率，其中有一个人的帽子不在这 $n-1$ 顶帽子中．此事件由两个互不相容的事件组成，一个事件是都不匹配且多余的那个人（即其帽子已被第一个人取走的那个人）未能选中多余的帽子（即第一个选取人的帽子）；另一个事件是都不匹配但多余的人刚好选到了多余的帽子．前一个事件的概率是 P_{n-1}，因为可以将多余的帽子看作为多余的人的，第二个事件的概率是 $\frac{1}{n-1}P_{n-2}$，所以有

$$P\{E \mid \overline{M}\} = P_{n-1} + \frac{1}{n-1}P_{n-2}$$

进而有

$$P_n = P\{E \mid \overline{M}\}\frac{n-1}{n} = \frac{n-1}{n}P_{n-1} + \frac{1}{n}P_{n-2}$$

或等价地

$$P_n - P_{n-1} = -\frac{1}{n}(P_{n-1} - P_{n-2})$$

显然有 $P_1 = 0, P_2 = \frac{1}{2}$，于是根据上式可得

$$P_3 - P_2 = -\frac{1}{3}(P_2 - P_1) = -\frac{1}{3!}$$

所以

$$P_3 = \frac{1}{2!} - \frac{1}{3!}$$

利用归纳法，可得 $P_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + \frac{(-1)^n}{n!}$（$n \geq 2$ 时）．

对于固定的 k 个人，只有他们选中自己帽子的概率为

$$\frac{1}{n}\cdot\frac{1}{n-1}\cdot\cdots\cdot\frac{1}{n-(k-1)}P_{n-k} = \frac{(n-k)!}{n!}P_{n-k}$$

其中 P_{n-k} 是其余 $n-k$ 个人从他们自己的那些帽子中选取但全都不匹配的概率．因 k 个人的选法有 C_n^k 种，所以恰好有 k 个匹配的概率是

$$\frac{(n-k)!}{n!}P_{n-k}\mathrm{C}_n^k = \frac{1}{k!}\left[\frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{n-k}}{(n-k)!}\right]$$

当 n 充分大时，上式近似为 $\mathrm{e}^{-1}/k!$．

【例 1.18】　设 $U \sim U[0,1]$，在 $U = p$ 的条件下，$X \sim B(n,p)$，试求 X 的分布．

解

$$P\{X = k\} = \int P\{X = k \mid U = p\}\mathrm{d}F_U(p)$$

$$= \int_0^1 \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}\mathrm{d}p$$

$$= \frac{n!}{k!(n-k)!} \frac{k!(n-k)!}{(n+1)!}$$

$$= \frac{1}{(n+1)} \quad (k = 0,1,\cdots,n)$$

即 X 是 $0,1,\cdots,n$ 上的均匀分布.

【例 1.19】 随机向量 (X,Y) 的联合概率密度为

$$f(x,y) = \begin{cases} \frac{1}{y}e^{-\frac{x}{y}}e^{-y}, & 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{其他} \end{cases}$$

试求 $E(X \mid Y)$.

解

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x = \begin{cases} \int_0^{+\infty} \frac{1}{y}e^{-\frac{x}{y}}e^{-y}\mathrm{d}x, & y > 0 \\ 0, & \text{其他} \end{cases} = \begin{cases} e^{-y}, & y > 0 \\ 0, & \text{其他} \end{cases}$$

当 $y > 0$ 时,

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{1}{y}e^{-\frac{x}{y}}, & x > 0 \\ 0, & \text{其他} \end{cases}$$

当 $y > 0$ 时,

$$E(X \mid Y = y) = \int_0^{+\infty} x \cdot f_{X|Y}(x \mid y)\mathrm{d}x = \int_0^{+\infty} x \cdot \frac{1}{y}e^{-\frac{x}{y}}\mathrm{d}x = y$$

所以

$$E(X \mid Y) = Y$$

【例 1.20】 设在底层乘坐电梯的人服从均值为 10 的 Poisson 分布,此楼共有 $N+1$ 层,每一个乘客在每一层楼要求停下来离开是等可能的,而且与其余乘客是否在这一层停下来离开是相互独立的.求所有乘客都走出电梯之前,该电梯所停的平均次数.

解 设 $X_i = \begin{cases} 1, & \text{电梯在第 } i+1 \text{ 层停} \\ 0, & \text{其他} \end{cases}$

则 $X = \sum_{i=1}^{N} X_i$ 表示电梯停的总次数.

设在底层乘坐电梯的人数为 $Y \sim P(10)$,则有

$$E(X) = \sum_{i=1}^{N} E(X_i) = \sum_{i=1}^{N} P\{X_i = 1\}$$

$$P\{X_i = 1\} = \sum_{k=1}^{+\infty} P\{Y = k\}P\{X_i = 1 \mid Y = k\}$$

$$= \sum_{k=1}^{+\infty} \frac{10^k}{k!}e^{-10}\left[1 - \left(1 - \frac{1}{N}\right)^k\right]$$

$$= \sum_{k=1}^{+\infty} \left[\frac{10^k}{k!} e^{-10} - \frac{\left[10\left(1 - \frac{1}{N}\right)\right]^k}{k!} e^{-10} \right]$$

$$= 1 - e^{-\frac{10}{N}}$$

所以

$$E(X) = N(1 - e^{-\frac{10}{N}})$$

1.6 特征函数

定义 1.15 随机变量 X 的分布函数为 $F(x)$，如果积分

$$\psi(t) \triangleq E(e^{tX}) = \int_{-\infty}^{+\infty} e^{tx} \, dF(x)$$

存在，则称其为 X 的矩母函数。

显然若 X 的 k 阶矩存在，则 $E(X^k) = \psi^{(k)}(0)$. 矩母函数由此得名.

当矩母函数存在时，它唯一地决定分布，因此可以用矩母函数来刻画随机变量的概率分布.

【例 1.21】 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$ 且相互独立，则它们的和 $Z = X + Y$ 的矩母函数为

$$\begin{aligned}\psi_Z(t) &= E[e^{t(X+Y)}] = E(e^{tX})E(e^{tY}) \\ &= \psi_X(t)\psi_Y(t) \\ &= \exp\{(\mu_1 + \mu_2)t + (\sigma_1^2 + \sigma_2^2)t^2/2\}\end{aligned}$$

由于 $X + Y$ 的矩母函数刚好是均值为 $\mu_1 + \mu_2$，方差为 $\sigma_1^2 + \sigma_2^2$ 的正态随机变量的矩母函数，根据唯一性知，$X + Y$ 的分布就是正态分布.

但随机变量的矩母函数不一定存在，因此在理论上，一般用如下定义的特征函数：

定义 1.16 设随机变量 X 的分布函数为 $F(x)$，称

$$\varphi(t) = E(e^{itx}) = \int_{-\infty}^{+\infty} e^{itx} \, dF(x), \quad -\infty < t < +\infty$$

为 X 的特征函数.

特征函数 $\varphi(t)$ 是实变量 t 的复值函数，因为 $|e^{itx}| = 1$，所以随机变量的特征函数一定存在. 显然特征函数只与分布函数有关，因此也称其为分布函数的特征函数.

若 X 为离散型随机变量，其分布列为 $p_k = P\{X = x_k\}$，$k = 1, 2, \cdots$，则其特征函数为

$$\varphi(t)=\sum_{k=1}^{+\infty}e^{itx_k}p_k,-\infty<t<+\infty$$

若 X 为连续型随机变量,其概率密度为 $f(x)$,则其特征函数为

$$\varphi(t)=\int_{-\infty}^{+\infty}e^{itx}f(x)dx,-\infty<t<+\infty$$

特征函数具有如下基本性质:

(1) 有界性:$|\varphi(t)|\leqslant|\varphi(0)|=1,-\infty<t<+\infty$.

(2) 一致连续性:$\varphi(t)$ 在整个实数轴$(-\infty,+\infty)$上一致连续,即对 $\forall\varepsilon>0$, $\exists\delta>0$,使得当$|h|<\delta$时,对 t 一致有

$$|\varphi(t+h)-\varphi(t)|<\varepsilon$$

(3) 非负定性:对于任意 $n\geqslant1$,任意实数 $t_1,t_2,\cdots,t_n$ 和任意复数 $z_1,z_2,\cdots,z_n$ 有

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\varphi(t_j-t_k)z_j\bar{z}_k\geqslant0$$

(4) 共轭对称性:$\varphi(-t)=\overline{\varphi(t)}$.

(5) 设 $Y=aX+b$,其中 a、b 为任意实数,$\varphi_X(t)$,$\varphi_Y(t)$ 分别是 X,Y 的特征函数,有

$$\varphi_Y(t)=a^{itb}\cdot\varphi_X(at)$$

(6) 设 X 和 Y 为独立随机变量,则

$$\varphi_{X+Y}(t)=\varphi_X(t)\varphi_Y(t)$$

一般,如果 $X_1,X_2,\cdots,X_n$ 相互独立,则

$$\varphi_{X_1+X_2+\cdots+X_n}(t)=\varphi_{X_1}(t)\varphi_{X_2}(t)\cdots\varphi_{X_n}(t)$$

(7) 如果 X 有 n 阶原点矩,则它的特征函数 $\varphi(t)$ 有 n 阶导数,且

$$E(X^k)=(-i)^k\varphi^{(k)}(0),k=1,2,\cdots,n$$

(8) 唯一性定理:若 X 的特征函数为 $\varphi(t)$,则 X 的分布函数 $F(x)$ 由 $\varphi(t)$ 唯一确定.

对于 n 维随机变量也可以定义特征函数如下:

定义 1.17　设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 是 n 维随机变量,$\boldsymbol{t}=(t_1,t_2,\cdots,t_n)\in\mathbf{R}^n$,称

$$\varphi_X(t)=\varphi_X(t_1,t_2,\cdots,t_n)=E[\exp(i\sum_{k=1}^{n}t_kX_k)]$$

为 X 的特征函数.

特征函数是研究概率论最重要的分析工具之一,用特征函数作为随机变量的研究工具比用分布函数要方便很多.例如,独立随机变量之和的概率分布是各被加项分布的卷积,而独立随机变量之和的特征函数则是各被加项特征函数的普通乘积.又如,通过随机变量的分布函数求其数字特征,一般需要进行积分运算,而通过

特征函数求其数字特征，一般只需要进行微商运算.

【例 1.22】 设随机变量 X 服从自由度为 n 的 χ^2 分布.

(1) 利用特征函数求 $E(X)$ 和 $D(X)$.

(2) 若 $X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，且 X,Y 相互独立，利用特征函数证明 $X+Y\sim\chi^2(n_1+n_2)$，即 χ^2 分布具有可加性.

解 (1) 因为 $X\sim\chi^2(n)$，故 X 的特征函数为 $\varphi_X(t)=(1-2\mathrm{i}t)^{-n/2}$，

$$\varphi'_X(0)=n\mathrm{i},$$
$$\varphi''_X(0)=-n(n+2),$$
$$E(X)=(-\mathrm{i})\varphi'_X(0)=n,$$
$$E(X^2)=(-\mathrm{i})^2\varphi''_X(0)=n(n+2),$$
$$D(X)=E(X^2)-[E(X)]^2=2n.$$

(2) 令 $Z=X+Y$，则

$$\begin{aligned}\varphi_Z(t)&=E[\mathrm{e}^{\mathrm{i}t(X+Y)}]=E(\mathrm{e}^{\mathrm{i}tX}\mathrm{e}^{\mathrm{i}tY})=\varphi_X(t)\varphi_Y(t)\\&=(1-2\mathrm{i}t)^{-n_1/2}\cdot(1-2\mathrm{i}t)^{-n_2/2}\\&=(1-2\mathrm{i}t)^{-\frac{n_1+n_2}{2}}\end{aligned}$$

即 $Z=X+Y\sim\chi^2(n_1+n_2)$

【例 1.23】 设 $X\sim N(0,1)$，ξ 与 X 相互独立，$P\{\xi=0\}=P\{\xi=1\}=\frac{1}{2}$，令

$$Y=\begin{cases}X, & \xi=0\\-X, & \xi=1\end{cases}$$

证明：$Y\sim N(0,1)$，但 (X,Y) 不服从二维正态分布.

证明 因为

$$\begin{aligned}\varphi_Y(t)&=E(\mathrm{e}^{\mathrm{i}tY})=E[E(\mathrm{e}^{\mathrm{i}tY}\mid\xi)]\\&=P\{\xi=0\}\cdot E(\mathrm{e}^{\mathrm{i}tY}\mid\xi=0)+P\{\xi=1\}\cdot E(\mathrm{e}^{\mathrm{i}tY}\mid\xi=1)\\&=\frac{1}{2}E(\mathrm{e}^{\mathrm{i}tX})+\frac{1}{2}E[\mathrm{e}^{\mathrm{i}t(-X)}]\\&=\mathrm{e}^{-\frac{1}{2}t^2}\end{aligned}$$

所以 $Y\sim N(0,1)$.

而

$$\begin{aligned}\varphi_{XY}(u,v)&=E(\mathrm{e}^{\mathrm{i}uX+\mathrm{i}vY})=E[E(\mathrm{e}^{\mathrm{i}uX+\mathrm{i}vY}\mid\xi)]\\&=P\{\xi=0\}\cdot E(\mathrm{e}^{\mathrm{i}uX+\mathrm{i}vY}\mid\xi=0)+P\{\xi=1\}\cdot E(\mathrm{e}^{\mathrm{i}uX+\mathrm{i}vY}\mid\xi=1)\\&=\frac{1}{2}E(\mathrm{e}^{\mathrm{i}uX+\mathrm{i}vX})+\frac{1}{2}E(\mathrm{e}^{\mathrm{i}uX-\mathrm{i}vX})\\&=\frac{1}{2}\mathrm{e}^{-\frac{1}{2}(u+v)^2}+\frac{1}{2}\mathrm{e}^{-\frac{1}{2}(u-v)^2}\end{aligned}$$

$$= \frac{1}{2}e^{-\frac{1}{2}(u^2+v^2)}(e^{-uv}+e^{uv})$$

故 $\varphi_{XY}(u,v)$ 不是二维正态分布的特征函数，因此 (X,Y) 不服从二维正态分布.

常用分布的特征函数见表 1.2.

表 1.2 常用分布的特征函数表

分布名称	分布律或概率密度	特征函数
单点分布	$p_a = 1$	e^{iat}
两点分布	$p_1 = p, p_0 = q, q = 1-p$	$pe^{it}+q$
二项分布 $B(n,p)$	$p_k = C_n^k p^k q^{n-k}, k=1,2,\cdots,n, q=1-p$	$(pe^{it}+q)^n$
泊松分布 $\pi(\lambda)$	$p_k = \frac{\lambda^k e^{-\lambda}}{k!}(\lambda>0), k=1,2,\cdots$	$e^{\lambda(e^{it}-1)}$
几何分布	$p_k = pq^{k-1}(0<p<1, q=1-p), k=1,2,\cdots$	$pe^{it}(1-qe^{it})^{-1}$
均匀分布	$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b \\ 0, & 其他\end{cases}$	$\frac{e^{ibt}-e^{iat}}{it(b-a)}$
指数分布	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0 \\ 0, & x\leqslant 0\end{cases}(\lambda>0)$	$(1-\frac{it}{\lambda})^{-1}$
正态分布 $N(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}(\sigma>0)$	$e^{i\mu t-\frac{\sigma^2t^2}{2}}$
$\chi^2(n)$ 分布	$f(x)=\begin{cases}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}}e^{-\frac{x}{2}}, & x>0 \\ 0, & x\leqslant 0\end{cases}$	$(1-2it)^{-\frac{n}{2}}$
$\Gamma(\alpha,\beta)$ 分布	$f(x)=\begin{cases}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}, & x>0 \\ 0, & x\leqslant 0\end{cases}$	$(1-\frac{it}{\beta})^{-\alpha}$

1.7 概率不等式

定理 1.11 (Chebyshev 不等式) 设随机变量 X 的期望为 $E(X)$，方差为 $D(X)$，则对任意给定的 $\varepsilon>0$，有

$$P\{|X-E(X)|\geqslant\varepsilon\}\leqslant\frac{D(X)}{\varepsilon^2}$$

或

$$P\{|X-E(X)|<\varepsilon\}\geqslant 1-\frac{D(X)}{\varepsilon^2}$$

Chebyshev 不等式只利用均值及方差就描述了随机变量的变化情况，不管 X 的分布是什么，X 落在 $(E(X)-\varepsilon, E(X)+\varepsilon)$ 中的概率都不小于 $1-\frac{D(X)}{\varepsilon^2}$. 因此它在理论研究及实际应用中很有价值.

【例 1.24】 森林中树木的高度 X 是随机变量，其均值 $\mu = 20$ m，标准差 $\sigma = 2$ m，问某棵树的高度与平均高度 20 m 至少相差 4 m 的概率有多大？

解 利用 Chebyshev 不等式知

$$P\{|X-20| \geqslant 4\} \leqslant 4/16 = 0.25$$

若 X 是正态随机变量，则

$$\begin{aligned} P\{|X-20| \geqslant 4\} &= P\{X-20 \geqslant 4\} + P\{X-20 \leqslant -4\} \\ &= 2\Phi(-2) = 0.046 \end{aligned}$$

其中 $\Phi(x)$ 是标准正态分布的分布函数.

此例也说明了 Chebyshev 不等式只给出了随机变量落在某个区间的一个粗略的估计.

定理 1.12 （Markov 不等式）若 $y = h(x)$ 是 $[0, +\infty)$ 上非负的严格递增的函数，则对任意 $\varepsilon > 0$，有

$$P\{|X| \geqslant \varepsilon\} \leqslant \frac{E[h(|X|)]}{h(\varepsilon)}$$

特别地，$h(x) = x^a$，$a > 0$ 时，Markov 不等式为

$$P\{|X| \geqslant \varepsilon\} \leqslant \frac{E(|X|^a)}{\varepsilon^a}$$

再令 $a = 2$，X 用 $X-\mu$ 替换则得到 Chebyshev 不等式.

定理 1.13 （Schwarz 不等式）对任意随机变量 X, Y，总是有

$$[E(|XY|)]^2 \leqslant E(X^2)E(Y^2)$$

等号成立当且仅当 $P\{Y = kX\} = 1$，k 是某一常数.

定理 1.14 （Holder 不等式）X, Y 是随机变量，正实数 r, s 满足 $\frac{1}{r} + \frac{1}{s} = 1$，则

$$E(|XY|) \leqslant [E(|X|^r)]^{1/r}[E(|Y|^s)]^{1/s}$$

令 $r = s = 2$，可得到 Schwarz 不等式.

定理 1.15 （Jensen 不等式）若 $h(x)$ 是一凸函数，则对任意随机变量 X，有

$$h[E(X)] \leqslant E[h(X)]$$

1.8 极限理论

定义 1.18 (1) 称$\{X_n\}$几乎处处收敛(或以概率为 1 收敛)于 X,记为 $X_n \xrightarrow{a.e} X$. 或 $\lim\limits_{n\to+\infty} X_n = X(a.e)$,如果

$$P\{\lim_{n\to+\infty} X_n = X\} = 1$$

(2) 称$\{X_n\}$依概率收敛于 X,记为 $X_n \xrightarrow{P} X$. 如果对 $\forall \varepsilon > 0$,有

$$\lim_{n\to+\infty} P\{|X_n - X| \geqslant \varepsilon\} = 0$$

(3) 设 X_n 的分布函数为 $F_n(x)$,X 的分布函数为 $F(x)$,如果对 $F(x)$ 的所有连续点 x,有

$$\lim_{n\to+\infty} F_n(x) = F(x)$$

则称$\{X_n\}$依分布收敛于 X,记为 $X_n \xrightarrow{d} X$.

(4) 设 $E|X_n|^r < +\infty, E|X|^r < +\infty, r \geqslant 1$,如果

$$\lim_{n\to+\infty} E|X_n - X|^r = 0$$

则称$\{X_n\}$$r$ 阶收敛于 X,记为 $X_n \xrightarrow{r} X$;当 $r = 2$ 时,称为均方收敛,记为 $\mathrm{l.i.m}X_n = X$.

这四种收敛性有下列关系:

$$X_n \xrightarrow{a.e} X. \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

$$X_n \xrightarrow{r} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

定理 1.16 (辛钦大数定律)$X_1, X_2, \cdots$ 是一列独立同分布的随机变量,且具有有限的数学期望 μ,则

$$\overline{X_n} = \frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{P} \mu$$

【例 1.25】 随机变量 X 表示随机事件 A 的示性函数,即 X 服从 Bernouli 分布,其中

$$X = \begin{cases} 1, & \text{若随机事件 } A \text{ 发生} \\ 0, & \text{若随机事件 } A \text{ 不发生} \end{cases}$$

$$p = P\{A\} = P\{X = 1\}, \quad 1 - p = P\{X = 0\} = P\{\overline{A}\}$$

为了估计 $p = P\{A\}$,重复独立地做 n 次试验,每次试验只有两个结果 A 和 $\overline{A}$,相应

结果的示性变量记为 $X_1, X_2, \cdots, X_n$，则 X_i 相互独立且与 X 同分布，所以有

$$\overline{X_n} = \frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{P} E(X) = p$$

注意 $\overline{X_n}$ 表示 n 次随机试验中随机事件 A 出现的频率，p 是随机事件 A 发生的概率，由上面的结论可知频率是概率的一个相合估计.

定理 1.17 （Kolmogorov 强大数定律）设 $X_1, X_2, \cdots$ 是相互独立的随机变量，期望为 $E(X_k)$，方差为 $D(X_k)$，如果 $\sum\limits_{k=1}^{+\infty} \frac{D(X_k)}{k^2} < +\infty$，则

$$P\{\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} [X_k - E(X_k)] = 0\} = 1$$

定理 1.18 （Linderberg-Levy 中心极限定理）设 $X_1, X_2, \cdots$ 是相互独立同分布的随机变量，具有均值 μ 和方差 σ^2，则

$$\lim_{n \to +\infty} P\{\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \leqslant a\} = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

定理 1.19 （Linderberg-Feller 中心极限定理）设 $X_1, X_2, \cdots$ 是一个相互独立的随机变量序列，它们具有有限的数学期望和方差：

$$\mu_i = E(X_i), \quad \sigma_i^2 = \mathrm{Var}(X_i)$$

记

$$Y_n = \sum_{i=1}^{n} X_n, \quad Z_n = \frac{Y_n - E(Y_n)}{\sqrt{\mathrm{Var}(Y_n)}} = \frac{Y_n - \sum\limits_{i=1}^{n} \mu_i}{\sqrt{\sum\limits_{i=1}^{n} \sigma_i^2}}$$

则

$$\lim_{n \to +\infty} P\{Z_n \leqslant x\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du, \lim_{n \to +\infty} E(Z_n) \to +\infty$$

和

$$\lim_{n \to +\infty} \max_{i=1,2,\cdots,n} \frac{\sigma_i}{E(Z_n)} \to 0$$

成立的充要条件是 Linderberg 条件

$$\lim_{n \to +\infty} \frac{1}{\mathrm{Var}(Z_n)} \int_{|x-\mu_i| > \varepsilon \sqrt{\mathrm{Var}(Z_n)}} (x - \mu_i)^2 dF_{X_i}(x) = 0$$

对所有 $\varepsilon > 0$ 都成立.

【例 1.26】 （正态分布逼近二项分布）记 $Y_n = X_1 + X_2 + \cdots + X_n$ 表示 n 次伯努利试验中事件 A 出现的次数，$Y_n \sim B(n, p)$，$E(Y_n) = np$，$\mathrm{Var}(Y_n) = np(1-p)$，根据 Linderberg-Levy 中心极限定理可知，Y_n 有近似正态分布 $N(np, np(1-p))$，因此

$$P\{i_1 \leqslant Y_n \leqslant i_2\} \approx \Phi\left(\frac{i_2+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{i_1-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right), 0 \leqslant i_1 \leqslant i_2 \leqslant n$$

特别地

$$P\{Y_n = i\} \approx \Phi\left(\frac{i+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{i-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right), 0 \leqslant i \leqslant n$$

上式中$\pm\frac{1}{2}$称为连续型修正，因为是用连续型分布逼近一个离散型分布，为了提高逼近的精度，加了这个连续型修正. 当n很大，p接近1/2时，一般$np > 35$，$np(1-p) > 10$，此时可以得到比较满意的正态逼近效果.

【例1.27】 已知在电子产品的质量检测中，有5%的产品是有故障的，问1000个电子产品中有故障的产品占的比率在4%～6%之间的概率有多大？

解 令X表示1000个产品中有故障的产品个数，则$X \sim B(1000, 0.05)$. 因此

$$P\{40 \leqslant X \leqslant 60\} = \sum_{i=40}^{60}\binom{1000}{i}(0.05)^i(0.95)^{1000-i}$$

为了简化计算，用二项分布的正态逼近求解.

又$np(1-p) = 1000 \times 0.05 \times 0.95 = 47.5 > 10$，可以得到比较满意的结果

$$P\{40 \leqslant X \leqslant 60\} \approx \Phi\left(\frac{60+0.5-50}{6.892}\right)-\Phi\left(\frac{40-0.5-50}{6.892}\right)$$
$$= \Phi(1.523)-\Phi(-1.523) = 0.972$$

第 2 章　随机过程的基本概念、类型和平稳随机过程

2.1　随机过程的概念

在概率论中，随机变量 X 是在固定条件下一个随机试验的结果，这些条件的变化会影响试验的结果，即 X 的概率分布会改变. 考虑到这些条件的变化，可以将随机变量 X 看作是依赖于一个确定性参数 t 的变量，即 $X = X(t)$. 这种定义的方法适用于广义的随机试验. 为了更好地描述这些广义的随机试验，先看下面的例子：

【例 2.1】　在一个特定的地理位置，每天 12:00 测量大气温度. x_i 表示在某年的第 i 天测量的温度，x_i 的值每天都在变化，因此，它可以看作是随机变量 X_i 的一次实现. 因此，在某年的第 i 天 12:00 测量的(随机)温度 X_i，除了温度的随机变化，X_i 也取决于一个确定性的变量，即时间. 如果仅对该年前三天(或任何其他连续的三天)的温度 X_1, X_2, X_3 感兴趣，那么这些温度至少是近似同分布的. 然而，如果要刻画出每天的温度之间明显存在的相依关系，则需要对随机向量(X_1, X_2, X_3)的联合概率分布有所了解.

从这个例子以及与之相联系的问题可以看出，有必要介绍下这个广义随机试验：在某年对一个给定的地理位置每天 12:00 的温度进行测量. 这个广义随机试验的随机结果是随机向量$(X_1, X_2, \cdots, X_{365})$，其中 X_i 既不独立也不同分布. 如果第 i 天的温度 x_i 测量出来了，那么向量$(x_1, x_2, \cdots, x_{365})$可以看作函数 $x = x(t), t \in \{1, 2, \cdots, 365\}$ 的值域：当 $t = i$ 时，$x(t) = x_i$. 向量$(x_1, x_2, \cdots, x_{365})$是随机向量$(X_1, X_2, \cdots, X_{365})$的一次实现.

如果一个传感器用图表记录了这一年的温度，那么测量的结果就是时间 t 的连续性函数：$x = x(t), 0 \leqslant t \leqslant 1$，其中 $x(t)$ 是随机温度 $X(t)$ 在一个固定地理位置上在时刻 t 处的一次实现. 因此，介绍这个广义随机试验——在一个给定的地理位置上一年内温度的连续测量是有意义的，它可以定义为$\{X(t), 0 \leqslant t \leqslant 1\}$.

这个广义随机试验的完整概率特性需要已知所有随机向量$(X(t_1), X(t_2), \cdots, X(t_n))$；$0 \leqslant t_1 < t_2 < \cdots < t_n \leqslant 1$；$n = 1, 2, \cdots$ 的联合概率分布. 很明显，对于很小的时间差异 $t_{i+1} - t_i$，$X(t_i)$ 和 $X(t_{i+1})$ 之间有一个很强的统计相依性. 然而，由于一

个地区天气形势的惯性，对于很长的时间差异 $t_k - t_i$，$X(t_i)$ 和 $X(t_k)$ 也可能是统计相关的.

【例 2.2】 影响随机试验结果确定性的变量不一定都是时间. 比如，在一个固定的时间点和固定的观察点，温度是根据距离地球表面的竖直长度 L 测量的，那么函数：$x = x(h), 0 \leqslant h \leqslant L$，明显依赖于测量点与地球表面的距离 h. 但是，如果试验在接下来的几年里在同样的条件（相同的时间、地点和测量程序）下重复进行，那么考虑到不可预测因素的影响，可以得到不同的函数 $x = x(h)$. 因此，在距离 h 下的温度是一个随机变量 $X(h)$. 这个广义随机试验 —— 根据距离地球表面的竖直长度 L 测量的温度，定义为：$\{X(h), 0 \leqslant h \leqslant L\}$. 它的一次实现是 h 的实函数：$x = x(h), 0 \leqslant h \leqslant L$.

在这个例子中，认为温度是距离 h 和观测的时间点 t 的函数也是有意义的，即：$x = x(h,t), 0 \leqslant h \leqslant L, t \geqslant 0$. 此时观测值 x 取决于一个确定性的随机向量：$\theta = (h,t)$. 不过，本书只考虑一维参数空间的情况.

随机试验的结果是具体数值，而广义随机试验的结果是具体函数，如例 2.1、例 2.2，因此，从字面上说，这种广义随机试验又经常被称作随机函数. 但是，术语"随机过程"更常用，本书沿用这个名称. 下面给出随机过程严格的数学定义：

定义 2.1 给定概率空间 $(\Omega, \mathscr{F}, P)$ 及参数集 T，对每一个参数 $t \in T$，$X(t, \omega)$ 是一定义在概率空间 $(\Omega, \mathscr{F}, P)$ 上的随机变量，则称随机变量族 $X_T = \{X(t, \omega), t \in T\}$ 为该概率空间上的随机过程，记为 $\{X(t), t \in T\}$ 或 $X(t)$.

参数集 T 一般表示时间或空间，常用的参数集有 $T = N_0 = \{0,1,2,\cdots\}$ 和 $T = [a,b]$，其中 a 可以取 0 或 $-\infty$，b 可以取 $+\infty$. 当参数集 T 取可列集时，一般称 $\{X(t), t \in T\}$ 为随机序列或时间序列，此时记为 $\{X_n, n \geqslant 1\}$；当参数集 T 是一个区间时，称 $\{X(t), t \in T\}$ 为连续时间随机过程.

$\{X(t), t \in T\}$ 的取值范围称为随机过程的状态空间，记作 I，I 中的元素称为状态. 若 I 为实数（复数）集，则称 $\{X(t), t \in T\}$ 为实（复）随机过程；若 I 为 n 维欧式空间，则称 $\{X(t), t \in T\}$ 为 n 维随机过程.

由定义知，随机过程 $\{X(t), t \in T\}$ 是一定义在 $T \times \Omega$ 上的二元函数，当固定 $t_0 \in T$ 时，$X(t_0, \cdot)$ 是定义在样本空间 Ω 上的一个随机变量；当固定 $\omega_0 \in \Omega$ 时，$X(\cdot, \omega_0)$ 是一个关于参数 $t \in T$ 的函数，通常称为样本函数（样本轨道），或称为随机过程的一次实现；当固定 $t_0 \in T, \omega_0 \in \Omega$ 时，$X(t_0, \omega_0) \in I$，即取到状态空间的一个元素.

【例 2.3】 连续不断地投掷一枚硬币，以"1"表示正面，"2"表示反面，$X_n(\omega)$ 表示第 n 次投掷的结果，则 $\{X_n, n \geqslant 1\}$ 是一随机过程.

【例 2.4】 考虑某"服务站"在 $[0,t]$ 内来的"顾客"数，记为 $N(t)$，固定 t，

$N(t)$ 就是一随机变量. 因此,$\{N(t), t \geqslant 0\}$ 是一随机过程. 这里的"顾客"可以是电话的"呼唤",通信设备中的"信号",一个系统的"更换设备",放射性物质衰变的"粒子"等.

【例 2.5】 在工程学、科学和经济学中,有许多依赖时间的随机现象,它们可以用随机过程来描述. 在一个电回路中,保持电压严格稳定是不可能的. 假设热噪声导致了电压的随机波动,如果 $v(t)$ 表示在时刻 t 电压的测量值,那么 $v = v(t)$ 就是随机过程$\{V(t), t \geqslant 0\}$ 的一条样本路径,其中 $V(t)$ 是在时刻 t 处电压的随机值.

设$\{X(t), t \in T\}$ 是一随机过程,依其定义可知,对任意固定的 $t \in T$,$X(t)$ 是一随机变量,其分布函数记为

$$F(t;x) = P\{X(t) \leqslant x\}, x \in \mathbf{R} \tag{2.1}$$

称式(2.1) 为随机过程$\{X(t), t \in T\}$ 的一维分布函数.

一般地,对 $\forall n \in \mathbf{N}, t_1, t_2, \cdots, t_n \in T$,$X(t_1), X(t_2), \cdots, X(t_n)$ 是 n 个随机变量,它们的 n 维联合分布函数为

$$\begin{aligned} &F(t_1, t_2, \cdots, t_n; x_1, x_2, \cdots, x_n) \\ &= P\{X(t_1) \leqslant x_1, X(t_2) \leqslant x_2, \cdots, X(t_n) \leqslant x_n\}, (x_1, x_2, \cdots, x_n) \in \mathbf{R}^n \end{aligned} \tag{2.2}$$

称式(2.2) 为随机过程$\{X(t), t \in T\}$ 的 n 维联合分布函数.

定义 2.2 设$\{X(t), t \in T\}$ 是一随机过程,其有限维联合分布函数的全体 $\{F(t_1, t_2, \cdots, t_n; x_1, x_2, \cdots, x_n), t_1, t_2, \cdots, t_n \in T, n \geqslant 1\}$称为随机过程$\{X(t), t \in T\}$ 的有限维分布函数族.

显然,随机过程的有限维分布函数族具有如下性质:

(1) 对称性:对$(1, 2, \cdots, n)$ 的任意排列$(i_1, i_2, \cdots, i_n)$,有

$$F(t_1, t_2, \cdots, t_n; x_1, x_2, \cdots, x_n) = F(t_{i_1}, t_{i_2}, \cdots, t_{i_n}; x_{i_1}, x_{i_2}, \cdots, x_{i_n})$$

(2) 相容性:对于 $m < n$,有

$$\begin{aligned} &F(t_1, t_2, \cdots, t_m, t_{m+1}, \cdots, t_n; x_1, x_2, \cdots, x_m, +\infty, \cdots, +\infty) \\ &= F(t_1, t_2, \cdots, t_m; x_1, x_2, \cdots, x_m) \end{aligned}$$

从理论上分析,对于给定的一个随机过程,可以先确定它的有限维分布函数族,并由此来研究随机过程的动态统计规律性. 但在一些实际问题中,并不是先定义出随机过程,而是先根据随机试验的结果,运用近似的方法提出有限维分布函数族,然后根据这个有限维分布函数族来建立一个随机过程的模型. 那么随机过程和有限维分布函数族是否存在一一对应的关系呢?Kolmogorov在 1931 年提出并解决了这个问题,下面不加证明地给出如下定理:

定理 2.1　（Kolmogorov 存在性定理）设分布函数族 $\{F(t_1,t_2,\cdots,t_n;x_1,x_2,\cdots,x_n),t_1,t_2,\cdots,t_n\in T,n\geqslant 1\}$ 满足上述的对称性和相容性，则必存在概率空间 $(\Omega,\mathscr{F},P)$ 及定义其上的随机过程 $\{X(t),t\in T\}$，使 $\{F(t_1,t_2,\cdots,t_n;x_1,x_2,\cdots,x_n),t_1,t_2,\cdots,t_n\in T,n\geqslant 1\}$ 恰好是 $\{X(t),t\in T\}$ 的有限维分布函数族.

2.2　随机过程的数字特征

Kolmogorov 存在性定理说明可以将对随机过程的研究转化为对有限维分布函数族的研究. 但是在实际问题中，要准确知道随机过程的全部有限维分布函数族是一件非常困难的事，有时甚至是不可能的. 因此，人们往往用随机过程的某些数字特征来刻画随机过程的局部性质，而且对于某些随机过程只需要求出其几个数字特征，就可以确定其有限维分布函数族.

2.2.1　均值函数

设 $\{X(t),t\in T\}$ 是随机过程，如果对任意 $t\in T$，$E[X(t)]$ 存在，则称函数

$$m_X(t)=E[X(t)]$$

为随机过程 $\{X(t),t\in T\}$ 的**均值函数**，$m_X(t)$ 表示随机过程在 t 时刻状态的统计平均.

如果对任意 $t\in T$，$E[X^2(t)]$ 存在，则称 $\{X(t),t\in T\}$ 为**二阶矩过程**.

2.2.2　（自）协方差函数

称

$$C_X(s,t)=E\{[X(s)-m_X(s)][X(t)-m_X(t)]\}$$

为 $\{X(t),t\in T\}$ 的（**自**）**协方差函数**，特别地，称

$$C_X(t,t)=\mathrm{Var}[X(t)]=E\{[X(t)-m_X(t)]^2\}$$

为 $\{X(t),t\in T\}$ 的**方差函数**，有时也记为 $\sigma_X^2(t)$ 或 $D_X(t)$，它反映了随机过程的样本函数在 t 时刻偏离 $m_X(t)$ 的程度.

由于协方差函数可以度量 $X(s)$，$X(t)$ 间的相依性，也许有人会认为以下等式成立：

$$\lim_{|t-s|\to+\infty}C(s,t)=0 \tag{2.3}$$

然而，后面的例 2.6 证明事实并非这样.

2.2.3 (自)相关函数

称

$$R_X(s,t) = E[X(s)X(t)]$$

为随机过程$\{X(t), t \in T\}$的(自)**相关函数**.

由Schwartz不等式知,二阶矩过程的协方差函数和相关函数存在,且有下列关系:

$$C_X(s,t) = R_X(s,t) - m_X(s) \cdot m_X(t)$$

在实际问题中,有时需要考虑两个随机过程的关系.此时,需要引入互协方差函数和互相关函数来刻画它们之间的线性关系.

定义 2.3 $\{X(t), t \in T\}$和$\{Y(t), t \in T\}$是两个二阶矩过程,则称

$$C_{XY}(s,t) = E\{[X(s) - \mu_X(s)][Y(t) - \mu_Y(t)]\}$$

为$\{X(t), t \in T\}$和$\{Y(t), t \in T\}$的**互协方差函数**,称

$$R_{XY}(s,t) = E[X(s)Y(t)]$$

为$\{X(t), t \in T\}$和$\{Y(t), t \in T\}$的**互相关函数**.

如果对任意的两个参数$s, t \in T$,有$C_{XY}(s,t) = 0$,则称$\{X(t), t \in T\}$和$\{Y(t), t \in T\}$是**互不相关**的.若$R_{XY}(s,t) = 0$,则称$\{X(t), t \in T\}$和$\{Y(t), t \in T\}$**正交**.

互协方差函数和互相关函数有以下的关系:

$$C_{XY}(s,t) = R_{XY}(s,t) - m_X(s) \cdot m_Y(t)$$

【例 2.6】 设$X(t) = A\cos(\omega t)$,其中A是一个非负随机变量,且$E(A) < +\infty$.试求该过程的均值函数和协方差函数.

解 均值函数为

$$m(t) = E(A)\cos(\omega t)$$

协方差函数为

$$\begin{aligned} C_X(s,t) &= E\{[A\cos(\omega s)][A\cos(\omega t)]\} - m(s)m(t) \\ &= \{E(A^2) - [E(A)]^2\}[\cos(\omega s)][\cos(\omega t)] \end{aligned}$$

因此,

$$C_X(s,t) = \mathrm{Var}(A)[\cos(\omega s)][\cos(\omega t)]$$

显然,这个过程并没有式(2.3)描述的性质.实际上,

$$\rho(s,t) = \frac{\mathrm{Cov}[X(s), X(t)]}{\sqrt{\mathrm{Var}[X(s)]}\sqrt{\mathrm{Var}[X(t)]}} = \frac{C_X(s,t)}{\sigma_X(s)\sigma_X(t)} \equiv 1$$

【例 2.7】 (脉冲编码调制)信号源分别以p和$1-p$的概率形成符号0和1,且形成0、1的过程是相互独立的.符号0表示在单位时间内没有发送脉冲;符号1表示在单位时间内发送一个恒定幅度的脉冲.信号源在过去就已经开始运行.以这

种方式产生的一个随机信号是由随机过程$\{X(t),t\in(-\infty,+\infty)\}$表示的，其中

$$X(t)=\sum_{-\infty}^{+\infty}A_nh(t-n),\quad n\leqslant t<n+1 \tag{2.4}$$

其中$A_n(n=0,\pm1,\pm2,\cdots)$是由以下式子定义的独立的随机变量：

$$A_n=\begin{cases}0, & \text{概率为 } p\\ a, & \text{概率为 } 1-p\end{cases}$$

且函数$h(t)$为

$$h(t)=\begin{cases}1, & \text{当 } 0\leqslant t<1\\ 0, & \text{否则}\end{cases}$$

所以，对于任意的t，有

$$X(t)=\begin{cases}0, & \text{概率为 } p\\ a, & \text{概率为 } 1-p\end{cases}$$

例如，图2.1中的样本路径是由以下信号的部分序列生成的：

$$\cdots 1\ 0\ 1\ 1\ 0\ 0\ 1\cdots$$

该过程具有恒定的均值函数

$$m(t)\equiv a\cdot p\{X(t)=a\}+0\cdot P\{X(t)=0\}=a(1-p)$$

对于$n\leqslant s,t<n+1;n=0,\pm1,\pm2,\cdots$，有

$$\begin{aligned}E[X(s)X(t)]&=E[X(s)X(t)\mid X(s)=a]\cdot P\{X(s)=a\}\\&\quad+E[X(s)X(t)\mid X(s)=0]\cdot P\{X(s)=0\}\\&=a^2(1-p)\end{aligned}$$

如果$m\leqslant s<m+1,n\leqslant t<n+1,m\neq n$，则$X(s)$与$X(t)$相互独立.

因此，随机过程$\{X(t),t\in(-\infty,+\infty)\}$的协方差函数为

$$C_X(s,t)=\begin{cases}a^2p(1-p), & \text{当 } n\leqslant s,t<n+1;n=0,\pm1,\pm2,\cdots\\ 0, & \text{否则}\end{cases}$$

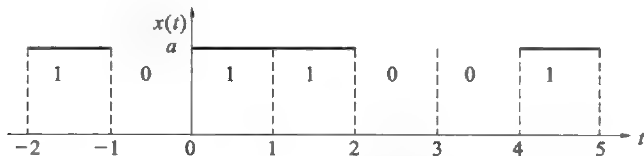


图2.1　脉冲编码调制

虽然这个例子中的随机过程是一个相当简单的结构，但它在物理、电气工程、通信方面是非常重要的.

【例2.8】　设$\{X_n,n\geqslant1\}$是独立同分布的随机变量序列，其中X_1的分布列为

X_1	1	-1
p	$\frac{1}{2}$	$\frac{1}{2}$

定义 $Y_n = \sum_{j=1}^{n} X_j$. 试对随机序列 $\{Y_n, n \geqslant 1\}$ 求：

(1)Y_1 的概率分布列；

(2)Y_2 的概率分布列；

(3)Y_n 的数学期望；

(4)Y_n 的(自)相关函数 $R_Y(m,n)$.

解 (1) 因为 $Y_1 = X_1$，故概率分布为 $P\{Y_1 = 1\} = \frac{1}{2}, P\{Y_1 = -1\} = \frac{1}{2}$.

(2) 因为 $Y_2 = X_1 + X_2$，Y_2 可能的取值为 0,2 或 -2.

$$P\{Y_2 = 0\} = P\{X_1 + X_2 = 0\} = P\{X_1 = 1, X_2 = -1\} + P\{X_1 = -1, X_2 = 1\}$$
$$= P\{X_1 = 1\}P\{X_2 = -1\} + P\{X_1 = -1\}P\{X_2 = 1\}$$
$$= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P\{Y_2 = 2\} = P\{X_1 + X_2 = 2\} = P\{X_1 = 1, X_2 = 1\} = \frac{1}{4}$$

$$P\{Y_2 = -2\} = P\{X_1 + X_2 = -2\} = P\{X_1 = -1, X_2 = -1\} = \frac{1}{4}$$

(3)$Y_n = \sum_{j=1}^{n} X_j$ 的数学期望为

$$E(Y_n) = E\left(\sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n} E(X_j) = \sum_{j=1}^{n}\left[1 \times \frac{1}{2} + (-1) \times \frac{1}{2}\right] = 0$$

(4)(自)相关函数为

$$R_Y(m,n) = E[Y(m)Y(n)] = E\left[\sum_{j=1}^{m} X_j \sum_{k=1}^{n} X_k\right]$$

当 $m \geqslant n$ 时，

$$R_Y(m,n) = E\left[\left(\sum_{j=1}^{n} X_j + \sum_{j=n+1}^{m} X_j\right)\sum_{k=1}^{n} X_k\right] = E\left[\left(\sum_{j=1}^{n} X_j\right)^2 + \sum_{j=n+1}^{m} X_j \sum_{k=1}^{n} X_k\right]$$
$$= E\left(\sum_{j=1}^{n} X_j\right)^2 + E\left(\sum_{j=n+1}^{m} X_j\right) \cdot E\left(\sum_{k=1}^{n} X_k\right)$$
$$= E(Y_n^2) + E\left(\sum_{j=n+1}^{m} X_j\right) \cdot E(Y_n) = E(Y_n^2)$$
$$= D(Y_n) + [E(Y_n)]^2 = D(Y_n)$$

因为 $$D(Y_n) = D\left(\sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n} D(X_j) \quad (X_j \text{ 相互独立})$$
$$= \sum_{j=1}^{n} \{[E(X_j)^2] - [E(X_j)]^2\}$$
$$E(X_j) = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0$$

$$E(X_j)^2 = 1$$

所以
$$D(Y_n) = \sum_{j=1}^{n}(1-0) = n$$

故当 $m \geqslant n$ 时，
$$R_Y(m,n) = D(Y_n) = n$$

同理当 $m < n$ 时，
$$R_Y(m,n) = D(Y_m) = m$$

所以
$$R_Y(m,n) = \min\{m,n\}$$

【例 2.9】 从 $t=0$ 开始每隔 1/2 s 丢掷一次硬币(均匀的)，对每一个丢掷的时刻 t，定义随机变量 $X(t) = \begin{cases} 2t, & \text{当时刻 } t \text{ 掷出正面} \\ \cos(\pi t), & \text{当时刻 } t \text{ 掷出反面} \end{cases}$.

试求：(1)$F\left(\frac{1}{2},x_1\right)$，$F(1,x_2)$；(2)$F\left(\frac{1}{2},1;x_1,x_2\right)$.

解　(1) 因为硬币是均匀的，所以“出现正面”和“出现反面”的概率各为$\frac{1}{2}$.

先求 $F\left(\frac{1}{2},x_1\right)$，显然

$$X\left(\frac{1}{2}\right) = \begin{cases} \cos\frac{\pi}{2}, & \text{出现反面} \\ 2\cdot\frac{1}{2}, & \text{出现正面} \end{cases} = \begin{cases} 0, & \text{出现反面} \\ 1, & \text{出现正面} \end{cases}$$

随机变量 $X\left(\frac{1}{2}\right)$的可能取值只有 0,1 两种可能，于是

$$P\left\{X\left(\frac{1}{2}\right)=0\right\} = \frac{1}{2}, \quad P\left\{X\left(\frac{1}{2}\right)=1\right\} = \frac{1}{2}$$

所以

$$F\left(\frac{1}{2},x_1\right) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \leqslant x < 1 \\ 1, & x \geqslant 1 \end{cases}$$

再求 $F(1,x_2)$，显然

$$X(1) = \begin{cases} \cos\pi, & \text{出现反面} \\ 2, & \text{出现正面} \end{cases} = \begin{cases} -1, & \text{出现反面} \\ 2, & \text{出现正面} \end{cases}$$

$$P\{X(1)=-1\} = P\{X(1)=2\} = \frac{1}{2}$$

所以

$$F(1,x_2) = \begin{cases} 0, & x < -1 \\ \frac{1}{2}, & -1 \leqslant x < 2 \\ 1, & x \geqslant 2 \end{cases}$$

(2) 计算 $F\left(\frac{1}{2},1;x_1,x_2\right)$

因为

$$X\left(\frac{1}{2}\right)=\begin{cases}0, & \text{出现反面}\\1, & \text{出现正面}\end{cases},\quad X(1)=\begin{cases}-1, & \text{出现反面}\\2, & \text{出现正面}\end{cases}$$

于是

$$F\left(\frac{1}{2},1;x_1,x_2\right)=P\left\{X\left(\frac{1}{2}\right)\leqslant x_1;X(1)\leqslant x_2\right\}$$

$$=\begin{cases}0, & x_1<0,-\infty<x_2<+\infty \text{ 或 } x_1\geqslant 0,\ x_2<-1\\ \frac{1}{2}, & 0\leqslant x_1<1,x_2\geqslant 2 \text{ 或 } x_1>1,\ -1\leqslant x_2<2\\ 1, & x_1>1,x_2\geqslant 2\end{cases}$$

【例 2.10】 给定随机过程 $\{X(t),-\infty<t<+\infty\}$,对于任意一个数 x,定义另一个随机过程

$$Y(t)=\begin{cases}1, & X(t)\leqslant x\\0, & X(t)>x\end{cases}$$

试证:$Y(t)$ 的数学期望和相关函数分别为随机过程 $X(t)$ 的一维和二维分布函数.

证明 设 $X(t)$ 的一维和二维概率密度分别为 $f_1(x,t)$ 和 $f_2(x_1,x_2;t_1,t_2)$,则

$$E[Y(t)]=\int_{-\infty}^{+\infty}y(t)f_1(x,t)\mathrm{d}x=\int_{-\infty}^{x}y(t)f_1(x,t)\mathrm{d}x+\int_{x}^{+\infty}y(t)f_1(x,t)\mathrm{d}t$$

$$=\int_{-\infty}^{x}f_1(x,t)\mathrm{d}t=F_1(x,t)$$

$$R_Y(t_1,t_2)=E[Y(t_1)Y(t_2)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}y_1y_2f_2(x_1,x_2;t_1,t_2)\mathrm{d}x_1\mathrm{d}x_2$$

$$=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}f_2(x_1,x_2;t_1,t_2)\mathrm{d}x_1\mathrm{d}x_2=F(x_1,x_2,t_1,t_2)$$

若考虑到对任意的 $t\in T$,$Y(t)$ 是离散型随机变量,则有

$$E_Y(t)=E[Y(t)]=1\cdot P\{Y(t)=1\}+0\cdot P\{Y(t)=0\}$$

$$=P\{X(t)\leqslant x\}=F_1(x,t)$$

$$R_Y(t_1,t_2)=E[Y(t_1)Y(t_2)]=1\cdot 1\cdot P\{Y(t_1)=1,Y(t_2)=1\}$$

$$+1\cdot 0\cdot P\{Y(t_1)=1,Y(t_2)=0\}+0\cdot 1\cdot P\{Y(t_1)=0,Y(t_2)=1\}$$

$$+0\cdot 0\cdot P\{Y(t_1)=0,Y(t_2)=0\}$$

$$=P\{X(t_1)\leqslant x_1,X(t_2)\leqslant x_2\}=F_2(x_1,x_2;t_1,t_2)$$

【例 2.11】 设随机过程 $\{X(t),-\infty<t<+\infty\}$ 共有三条样本函数:

$$X(t,\bar{\omega}_1)=1,\quad X(t,\bar{\omega}_2)=\sin t,\quad X(t,\bar{\omega}_3)=\cos t$$

且 $P\{\{\bar{\omega}_1\}\}=P\{\{\bar{\omega}_2\}\}=P\{\{\bar{\omega}_3\}\}=\frac{1}{3}$,试求随机过程 $X(t)$ 的均值函数 $\mu_X(t)$ 和

相关函数 $R_X(t_1,t_2)$.

解　$X(t)$的均值函数为

$$\mu_X(t)=E[X(t)]=1\cdot\frac{1}{3}+\sin t\cdot\frac{1}{3}+\cos t\cdot\frac{1}{3}=\frac{1}{3}+\frac{1}{3}(\sin t+\cos t).$$
$$=\frac{1}{3}(1+\sin t+\cos t)$$

相关函数为

$$R_X(t_1,t_2)=E[X(t_1)X(t_2)]=1\cdot\frac{1}{3}+\sin t_1\cdot\sin t_2\cdot\frac{1}{3}+\frac{1}{3}\cos t_1\cos t_2$$
$$=\frac{1}{3}[1+\cos(t_1-t_2)]$$

【例 2.12】　设随机过程 $X(t)=X+Yt$, $-\infty<t<+\infty$,随机向量(X,Y)的协方差矩阵为$\begin{bmatrix}\sigma_1^2 & \gamma\\ \gamma & \sigma_2^2\end{bmatrix}$,试求 $X(t)$ 的协方差函数.

解　因为$C_X(t_1,t_2)=E[X(t_1)X(t_2)]-E[X(t_1)]\cdot E[X(t_2)]$

而　$X(t_1)X(t_2)=(X+Yt_1)(X+Yt_2)=X^2+XYt_2+XYt_1+Y^2t_1t_2$

则
$$E[X(t_1)X(t_2)]=E[X^2+XYt_2+XYt_1+Y^2t_1t_2]$$
$$=E(X^2)+t_1t_2E(Y^2)+(t_1+t_2)E(XY)$$
$$E(X+Yt_1)\cdot E(X+Yt_2)=[E(X)+t_1E(Y)][E(X)+t_2E(Y)]$$
$$=[E(X)]^2+t_2E(X)\cdot E(Y)+t_1E(X)\cdot E(Y)+t_1t_2[E(Y)]^2$$
$$=[E(X)]^2+(t_1+t_2)E(X)\cdot E(Y)+t_1t_2[E(Y)]^2$$

则
$$C_X(t_1,t_2)=\{E(X^2)-[E(X)]^2\}+t_1t_2\{E(Y^2)-[E(Y)]^2\}+(t_1+t_2)[E(XY)-E(X)E(Y)]$$
$$=D(X)+t_1t_2D(Y)+(t_1+t_2)\mathrm{Cov}(X,Y)$$
$$=\sigma_1^2+t_1t_2\sigma_2^2+(t_1+t_2)\gamma$$

2.3　随机过程的分类

根据随机过程的不同性质,比如是否依赖于时间,不相交的时间段内过程是否是独立的,过程现在或过去的状态对未来状态的影响等,可以将随机过程分为强平稳过程、弱平稳过程、平稳增量过程、独立增量过程、高斯过程、Markov 过程等类型.

2.3.1 强平稳过程

称随机过程$\{X(t),t\in T\}$是**强平稳或严平稳过程**，如果对于任意的$t_1,t_2,\cdots,t_n\in T$和任意的h，且$t_i+h\in T$，有

$$(X(t_1),X(t_2),\cdots,X(t_n))\stackrel{d}{=}(X(t_1+h),X(t_2+h),\cdots,X(t_n+h))\tag{2.5}$$

这里“d”表示等式两边n维随机向量的分布相同.

由定义可知，强平稳过程的概率分布关于时间是平移不变的. 特别地，当$n=1$时，式(2.5)表明一维分布$F_t(x)$不依赖于t，在这种情况下，存在一维分布函数$F(x)$，使得

$$F_t(x)\equiv F(x),t\in T\tag{2.6}$$

因此，均值函数和方差函数也不依赖于t，即有

$$m(t)=E[X(t)]\equiv m=\text{常数}\tag{2.7}$$

$$\mathrm{Var}[X(t)]\equiv\text{常数}\tag{2.8}$$

2.3.2 弱平稳过程

在实际情况下，不可能为了验证一随机过程是否是强平稳过程，而去确定随机向量组$\{X(t_1),X(t_2),\cdots,X(t_n)\}$的联合概率分布函数. 因此，现给出条件稍弱的弱平稳过程：

如果随机过程$\{X(t),t\in T\}$是二阶过程，并且满足

(1)$E[X(t)]=m_X=$常数；

(2)$R_X(t_1,t_2)=E[X(t_1)X(t_2)]=R_X(t_1-t_2)$.

则称$\{X(t);t\in T\}$为**宽(弱)平稳过程**.

由于有的强平稳过程不是二阶过程，所以相对于弱平稳而言，强平稳是不必要的. 但是如果二阶过程是强平稳过程，那么它一定是弱平稳过程. 弱平稳过程也被称为广义平稳，协方差平稳或二阶平稳随机过程. 弱平稳过程更重要的性质是它们的协方差函数、相关函数只与时间差有关.

本书以后凡提到“平稳过程”，若不加说明，通常是指宽(弱)平稳过程.

2.3.3 平稳增量过程

记随机过程$\{X(t),t\in T\}$在区间$[t_1,t_2]$上的增量为$X(t_2)-X(t_1)$.

如果对任意的$t_1,t_2\in T,\tau+t_1\in T,\tau+t_2\in T$，有$X(t_1+\tau)-X(t_1)\stackrel{d}{=}X(t_2+\tau)-X(t_2)$，则称过程$\{X(t);t\in T\}$为**平稳增量过程**.

由此可知，对于平稳增量过程，增量 $X(t+\tau)-X(t)$ 的概率分布不依赖于 t，而只与时间差 τ 有关.

2.3.4 独立增量过程

如果对任意的 $n=1,2,3,\cdots$，任意的 $t_1,t_2,\cdots,t_n\in T,t_1<t_2<\cdots<t_n$，随机变量 $X(t_2)-X(t_1),X(t_3)-X(t_2),\cdots,X(t_n)-X(t_{n-1})$ 是相互独立的，则称 $\{X(t),t\in T\}$ 为**独立增量过程**.

2.3.5 高斯过程

称随机过程 $\{X(t),t\in T\}$ 为**高斯过程或正态过程**，如果对任意 $t_1,t_2,\cdots,t_n\in T,n=1,2,3,\cdots,(X(t_1),X(t_2),\cdots,X(t_n))$ 是 n 维正态随机向量. 高斯过程的一个重要的性质：高斯过程是强平稳过程，当且仅当它是弱平稳过程.

2.3.6 Markov 过程

随机过程 $\{X(t),t\in T\}$，如果对 $t_1,t_2,\cdots,t_{n+1}\in T,t_1<t_2<\cdots<t_{n+1}$，且 $\forall A_i\subseteq \mathbf{Z},i=1,2,\cdots,n+1$，有

$$
\begin{aligned}
&P\{X(t_{n+1})\in A_{n+1}\mid X(t_n)\in A_n,X(t_{n-1})\in A_{n-1},\cdots,X(t_1)\in A_1\}\\
&=P\{X(t_{n+1})\in A_{n+1}\mid X(t_n)\in A_n\} \qquad (2.9)
\end{aligned}
$$

则称 $\{X(t),t\in T\}$ 具有 Markov 性，具有 Markov 性的随机过程称为 **Markov 过程**.

由 Markov 性可知：如果 t_{n+1} 是未来的某个时间点，t_n 是当前时刻的时间点，相应地，$t_1,t_2,\cdots,t_{n-1}$ 是过去的时间点，一个具有 Markov 性的过程，其未来的发展趋势不依赖于过去的演变，而只与当前的状态有关.

当参数集 T 是有限集或可数集时，Markov 过程称为离散时间 Markov 过程，否则称为连续时间 Markov 过程. 当状态空间 I 是有限集或可数集时，Markov 过程称为 Markov 链. 因此，一个离散时间 Markov 链具有离散状态空间和离散参数空间.

$\{X(t),t\in T\}$ 是二阶过程，如果

$$
\lim_{h\to 0}E\{[X(t_0+h)-X(t_0)]^2\}=0,\ t_0\in T \qquad (2.10)
$$

则称 $\{X(t),t\in T\}$ 在点 t_0 处均方连续.

如果 $\{X(t),t\in T\}$ 在任意 $t\in T_0$ 处都均方连续，则称过程在 $T_0(T_0\subseteq T)$ 内均方连续.

【例 2.13】 设 Y 是随机变量，试分别考虑随机过程 $X_1(t)=Y,X_2(t)=tY$ 的

平稳性.

解 因为 Y 是随机变量,所以 $X_1(t)=Y$ 这一过程是一个与时间无关的特殊的过程,它的任何 n 维分布函数 $F_Y(y_1,y_2,\cdots,y_n)$ 都与时间无关,所以是一个严平稳.

因为 $X_1(t)=Y$ 是严平稳,所以只要 $E[X_1^2(t)]=E(Y^2)<+\infty$,即可得 $X_1(t)$ 是宽平稳.

对于 $X_2(t)=tY$,

$$E[X_2(t)]=E(tY)=tE(Y)$$

$$R_{X_2}(t_1,t_2)=E[X_2(t_1)X_2(t_2)]=E(t_1Yt_2Y)=t_1t_2E(Y^2)$$

与时间 t_1,t_2 有关,所以 $X_2(t)=tY$ 为非平稳.

【例 2.14】 $\{X(n),n=0,\pm1,\cdots\}$ 是实的、互不相关的随机变量序列,且满足 $E[X(n)]=0,D[X(n)]=\sigma^2$. 由于

$$R_X(m,n)=E[X(m)X(n)]=\begin{cases}\sigma^2, & m=n\\ 0, & m\neq n\end{cases}$$

只与时间差 $m-n$ 有关,均值为常数,因此 $\{X(n),n=0,\pm1,\cdots\}$ 是平稳时间序列.

在科学和工程技术中,上述的时间序列称为“白噪声”.

【例 2.15】 $\{X(n),n=0,\pm1,\cdots\}$ 是白噪声,设 $a_0,a_1,\cdots,a_k$ 为任意实数,令 $Y(n)=a_0X(n)+a_1X(n-1)+\cdots+a_kX(n-k)$,证明:$\{Y(n),n=0,\pm1,\cdots\}$ 是一平稳时间序列.

证明 显然 $E[Y(n)]=0$

$$\begin{aligned}R_Y(\tau)&=E[Y(n+\tau)Y(n)]\\&=E[\sum_{i=0}^{k}a_iX(n+\tau-i)\sum_{j=0}^{k}a_jX(n-j)]\\&=\sum_{i=0}^{k}\sum_{j=0}^{k}a_ia_jE[X(n+\tau-i)X(n-j)]\\&=\sum_{0\leqslant i\leqslant k,0\leqslant i-\tau\leqslant k}a_ia_{i-\tau}\\&=\begin{cases}a_ka_{k-\tau}+\cdots+a_\tau a_0, & \text{当 } 0\leqslant\tau\leqslant k\\ 0, & \text{当 } \tau\geqslant k+1\end{cases}\end{aligned}$$

只与时间差 τ 有关,故 $\{Y(n),n=0,\pm1,\cdots\}$ 是一平稳时间序列. 形如上式的 $\{Y(n),n=0,\pm1,\cdots\}$ 又称为滑动平均序列.

下面的两个例子用到了三角公式中的两个加法公式:

$$\cos\alpha\cos\beta=\frac{1}{2}[\cos(\beta-\alpha)+\cos(\alpha+\beta)]$$

$$\cos(\beta-\alpha)=\cos\alpha\cos\beta+\sin\alpha\sin\beta$$

【例 2.16】 设随机过程 $X(t)=A\cos(\omega t+\Phi),-\infty<t<+\infty$,其中 A 是一个有有限均值和有限方差的非负随机变量,而随机变量 Φ 服从区间$[0,2\pi]$上的均匀分

布，且与 A 相互独立. 试证：随机过程 $\{X(t), t\in(-\infty,+\infty)\}$ 是弱平稳过程.

证明　$$E[\cos(\omega t+\Phi)]=\frac{1}{2\pi}\int_0^{2\pi}\cos(\omega t+\varphi)\mathrm{d}\varphi=\frac{1}{2\pi}[\sin(\omega t+\varphi)]\big|_0^{2\pi}=0$$

所以均值函数为

$$m(t)=E[X(t)]=E(A)E[\cos(\omega t+\Phi)]=0$$

协方差函数为

$$\begin{aligned}C_X(s,t)&=E\{[A\cos(\omega s+\Phi)][A\cos(\omega t+\Phi)]\}\\&=E(A^2)\frac{1}{2\pi}\int_0^{2\pi}\cos(\omega s+\varphi)\cos(\omega t+\varphi)\mathrm{d}\varphi\\&=E(A^2)\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}\{\cos[\omega(t-s)]+\cos[\omega(s+t)+2\varphi]\}\mathrm{d}\varphi\\&=\frac{1}{2}E(A^2)\cos[\omega(t-s)]\end{aligned}$$

并且　$$E[X^2(t)]=C_X(t,t)=\frac{1}{2}A^2<+\infty$$

因此，此过程是弱平稳的.

【例 2.17】　给定随机过程 $X(t)=A\cos(\omega t)+B\sin(\omega t)$，式中 ω 是常数，A 和 B 是两个独立的正态随机变量，而且 $E(A)=E(B)=0, E(A^2)=E(B^2)=\sigma^2$，试证：$X(t)$ 是弱平稳过程.

解　因为 $X(t)=A\cos(\omega t)+B\sin(\omega t)$，且 A,B 相互独立

所以　$$E(AB)=E(A)\cdot E(B)=0$$

$$\begin{aligned}E[X(t)]&=E[A\cos(\omega t)]+E[B\sin(\omega t)]\\&=\cos(\omega t)E(A)+\sin(\omega t)E(B)=0\end{aligned}$$

$$\begin{aligned}C_X(t_1,t_2)&=E[X(t_1)X(t_2)]-0\\&=E\{[A\cos(\omega t_1)+B\sin(\omega t_1)]\cdot[A\cos(\omega t_2)+B\sin(\omega t_2)]\}\\&=E[A^2\cos(\omega t_1)\cos(\omega t_2)+AB\cos(\omega t_1)\sin(\omega t_2)\\&\quad+BA\sin(\omega t_1)\cos(\omega t_2)+B^2\sin(\omega t_1)\sin(\omega t_2)]\\&=\cos(\omega t_1)\cos(\omega t_2)E(A^2)+\cos(\omega t_1)\sin(\omega t_2)E(AB)\\&\quad+\sin(\omega t_1)\cos(\omega t_2)E(AB)+\sin(\omega t_1)\sin(\omega t_2)E(B^2)\\&=\sigma^2\cos(\omega t_1)\cos(\omega t_2)+\sigma^2\sin(\omega t_1)\sin(\omega t_2)\\&=\sigma^2\cos[\omega(t_1-t_2)]\end{aligned}$$

并且　$$E[X^2(t)]=C_X(t,t)=\sigma^2<+\infty$$

所以，此过程是弱平稳的.

2.4　平稳随机过程的遍历性

考虑如下两个平稳过程：

$$X = \{X_n, n = 0,1,2,\cdots\}$$

其中 X_n 为独立同分布随机变量序列，$E(X_n^2) < +\infty$，$E(X_n) = m$，$n = 0,1,2,\cdots$；

$$Y = \{Y_n = Y, n = 0,1,2,\cdots\}$$

其中 Y 是随机变量，$E(Y^2) < +\infty$.

这两个随机过程是平稳过程中的两个极端情况，可以用这两个过程来阐述不同平稳过程之间的差异.

由大数定律知，对 X 而言，有

$$\frac{1}{n}(X_0 + X_1 + \cdots + X_{n-1})$$

以概率收敛于常数 m；但对 Y 过程来说，

$$\frac{1}{n}(Y_0 + Y_1 + \cdots + Y_{n-1}) = Y$$

即经过对时间的平均后，随机性没有什么改变. 于是自然产生这样的问题：对平稳过程加上什么条件后，对时间的平均值可以等于过程的均值？这一问题称为平稳过程的遍历性问题，这是平稳过程研究中的一个重要内容，其重要性可从如下粗略的分析中看出：

对于随机过程 $\{X(t), -\infty < t < +\infty\}$，由于 $E[X(t)] = m$，为估计 m，就必须对 $\{X(t), -\infty < t < +\infty\}$ 做大量的观测. 记 $X_i(t)$ 为第 i 次观测中时刻 t 的值，由大数定律知，可用 $\frac{1}{N}\sum_{i=1}^{N} X_i(t)$ 来估计 m. 同样为了估计相关函数 $R_X(\tau)$，可用 $\frac{1}{N}\sum_{i=1}^{N} X_i(t)X_i(t+\tau)$ 来估计.

但是这样的计算需要对一个随机过程重复地进行大量观察，以便获取很多样本函数 $X_i(t)$，而这在实际中是非常困难的，于是自然希望在很长时间内观察得到一个样本函数，由这一次观察来估计均值和相关函数. 对于一般的随机过程这是不可能的，但根据平稳过程不随时间的推移而变化的统计特性则是可能的，只要加上一些条件，就可以从一次观察中得到 m 和 $R_X(\tau)$ 的较好的估计.

根据随机过程的定义，对固定的 $t \in T$，$X(t)$ 是一定义在样本空间 Ω 上的函数，即为一随机变量，$E[X(t)] = m_X(t)$ 为统计平均；对于固定的 $\omega \in \Omega$，$X(\cdot, \omega)$ 是一个关于参数 $t \in T$ 的函数，若在 T 上取平均，则为时间平均.

本节给出的遍历性定理指出：对平稳过程而言，只要满足一些较宽的条件，那

么均值函数和相关函数实际上可以用一个样本函数在整个时间轴上的时间平均值来代替，也就是用随机过程的时间平均来代替其统计平均.

定义 2.4　设$\{X(t), -\infty < t < +\infty\}$是平稳随机过程，如果

$$\overline{X} = \underset{T\to+\infty}{\text{l.i.m}} \frac{1}{2T}\int_{-T}^{T} X(t)\mathrm{d}t = m_X \tag{2.11}$$

即

$$\lim_{T\to+\infty} E\left[\left|\frac{1}{2T}\int_{-T}^{T} X(t)\mathrm{d}t - m_X\right|^2\right] = 0$$

则称$\{X(t), -\infty < t < +\infty\}$的均值具有遍历性.

定义 2.5　设$\{X(t), -\infty < t < +\infty\}$是平稳随机过程，如果

$$\overline{R}(\tau) = \underset{T\to+\infty}{\text{l.i.m}} \frac{1}{2T}\int_{-T}^{T} X(t+\tau)X(t)\mathrm{d}t = R(\tau) \tag{2.12}$$

即

$$\lim_{T\to+\infty} E\left[\left|\frac{1}{2T}\int_{-T}^{T} X(t+\tau)X(t)\mathrm{d}t - R(\tau)\right|^2\right] = 0$$

则称$\{X(t), -\infty < t < +\infty\}$的相关函数具有遍历性.

定义 2.6　如果随机过程$\{X(t), -\infty < t < +\infty\}$的均值函数和相关函数均具有遍历性，则称该随机过程具有遍历性，或者说是各态历经的.

遍历性直观上可以这样理解：考虑只有有限个状态的平稳序列$\{X_n, n = 0, \pm 1, \pm 2, \cdots\}$，其状态空间$I = \{e_1, e_2, \cdots, e_b\}$，则$m = E(X_n)$是各个状态的加权平均. 若令$A_N = \{X_n, -N \leqslant n \leqslant N\}$，则由遍历性可知，对几乎每个样本，当$N$很大时，$A_N$中的元素历经$I$中各个状态，而且当$N \to +\infty$时，$A_N$中的元素为状态$e_i$的频率趋于$p_i$，从而$A_N$中元素的平均$\overline{X_N} = \frac{1}{2N+1}\sum_{n=-N}^{N} X_n$趋于过程均值$m = \sum_{i=1}^{b} p_i e_i$.

关于定义 2.4，定义 2.5 中的式子，当$X(t)$只取非负实数或整数时，相应的积分变为

$$\overline{X} = \underset{T\to+\infty}{\text{l.i.m}} \frac{1}{T}\int_{0}^{T} X(t)\mathrm{d}t = m_X$$

或

$$\overline{X} = \underset{N\to+\infty}{\text{l.i.m}} \frac{1}{2T+1}\sum_{k=-N}^{N} X(k) = m_X$$

【例 2.18】　设随机过程$\{X(t) = A\cos(\omega t + \varphi), -\infty < t < +\infty\}$，其中$A, \omega, \varphi$是相互独立的随机变量，$E(A) = 2, D(A) = 4, \omega \sim U(-5, 5), \varphi \sim U(-\pi, \pi)$，讨论$\{X(t), -\infty < t < +\infty\}$的遍历性.

解 (1)

$$m_X(t)=E[X(t)]=E[A\cos(\omega t+\varphi)]$$
$$=E(A)E[\cos(\omega t+\varphi)]$$
$$=2\cdot\frac{1}{20\pi}\cdot\int_{-5}^{5}\int_{-\pi}^{\pi}\cos(\omega t+\varphi)\mathrm{d}\omega\mathrm{d}\varphi=0$$

(2)

$$R_X(t,t+\tau)=E[X(t)X(t+\tau)]$$
$$=E[A\cos(\omega t+\varphi)A\cos(\omega t+\omega\tau+\varphi)]$$
$$=\frac{1}{2}E(A^2)E[\cos(\omega\tau)+\cos(2\omega t+\omega\tau+2\varphi)]$$
$$=8\cdot\frac{1}{2}\cdot\frac{1}{10}\int_{-5}^{5}\cos(\omega\tau)\mathrm{d}\omega$$
$$=\frac{4\sin(5\tau)}{5\tau}=R_X(\tau)$$

(3) $E[X^2(t)]=R_X(t,t)=R_X(0)=4<+\infty$,

由(1),(2),(3) 知$\{X(t),-\infty<t<+\infty\}$是平稳过程，故

$$\overline{X(t)}=\underset{T\to+\infty}{\mathrm{l.i.m}}\frac{1}{2T}\int_{-T}^{T}A\cos(\omega t+\varphi)\mathrm{d}t$$
$$=\underset{T\to+\infty}{\mathrm{l.i.m}}\frac{A}{\omega T}\sin(\omega T)\cos\varphi$$
$$=0=m_X$$

$$\overline{R}(\tau)=\underset{T\to+\infty}{\mathrm{l.i.m}}\frac{1}{2T}\int_{-T}^{T}A\cos(\omega t+\varphi)\cdot A\cos(\omega t+\omega\tau+\varphi)\mathrm{d}t$$
$$=\underset{T\to+\infty}{\mathrm{l.i.m}}\frac{A^2}{2T}\int_{-T}^{T}\cos(\omega t+\varphi)\cdot\cos(\omega t+\omega\tau+\varphi)\mathrm{d}t$$
$$=\frac{A^2}{2}\cos(\omega\tau)\neq R_X(\tau)$$

因此，$\{X(t),-\infty<t<+\infty\}$的均值具有遍历性，但相关函数不具有遍历性.

由遍历性的定义，很自然要问式(2.10)、式(2.11)左边极限是否存在？如果极限存在，而且等式左边是随机变量的均方极限，那么自然要问在什么条件下它能等于常数 m？关于第一个问题，主要是数学上的考虑. 1931 年 Birkhoff 证明了只要 $E[|X(t)|]<+\infty$，则对几乎所有的样本，式(2.10)左边的极限一定存在. 第二个问题就是本节所要解决的问题.

定理 2.2 （均值遍历性定理）

(1) 平稳过程$\{X(t),-\infty<t<+\infty\}$的均值具有遍历性的充要条件是

$$\lim_{T\to+\infty}\frac{1}{T}\int_{0}^{2T}\left(1-\frac{\tau}{2T}\right)C_X(\tau)\mathrm{d}\tau=0$$

(2) 平稳序列$\{X_n, n=0,\pm1,\pm2,\cdots\}$的均值具有遍历性的充要条件是

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{\tau=0}^{N-1}C_X(\tau)=0$$

式中$C_X(\tau)$为$\{X(t), -\infty<t<+\infty\}$的协方差函数.

证明

由于离散场合和连续时间场合证明思路相同，所以现仅证明连续时间的均值遍历性定理. 首先计算$\overline{X(t)}$的均值和方差，记

$$\overline{X}_T=\frac{1}{2T}\int_{-T}^{T}X(t)\mathrm{d}t$$

则

$$E(\overline{X}_T)=\frac{1}{2T}\int_{-T}^{T}E[X(t)]\mathrm{d}t=m$$

下面计算$\overline{X(t)}$的方差，

$$\begin{aligned}\mathrm{Var}(\overline{X}_T)&=E\left\{\frac{1}{2T}\int_{-T}^{T}X(t)\mathrm{d}t\right\}^2-\frac{1}{4T^2}\left\{\int_{-T}^{T}E[X(t)]\mathrm{d}t\right\}^2\\&=\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T}E\{[X(t)-m][X(s)-m]\}\mathrm{d}t\mathrm{d}s\\&=\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T}C_X(t-s)\mathrm{d}t\mathrm{d}s\end{aligned}$$

在上述积分中，令$\tau=t-s, v=t+s$，则变换的Jacobi行列式值为$\frac{1}{2}$，积分区域变换为$D=\{-2T\leqslant\tau\pm v\leqslant 2T\}$，注意到$C_X(\tau)$是偶函数，故上式可简化为

$$\begin{aligned}\frac{1}{4T^2}\iint_D C_X(\tau)\mathrm{d}\tau&=\frac{1}{4T^2}\int_{-2T}^{2T}C_X(\tau)(2T-|\tau|)\mathrm{d}\tau\\&=\frac{1}{2T^2}\int_0^{2T}C_X(\tau)(2T-\tau)\mathrm{d}\tau=\frac{1}{T}\int_0^{2T}C_X(\tau)(1-\frac{\tau}{2T})\mathrm{d}\tau\\&=\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T}C_X(t-s)\mathrm{d}t\mathrm{d}s\end{aligned}$$

故关于均值遍历性定理就化为上式极限是否趋于零的问题，由均方收敛定义知这确实是等价的，从而证明了定理.

【例 2.19】 已知随机电报信号$X(t)$，它的$E[X(t)]=0, R_X(\tau)=\mathrm{e}^{-\alpha|\tau|}$，问$X(t)$的均值是否各态历经.

解

$$\begin{aligned}&\lim_{T\to+\infty}\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau}{2T}\right)C_X(\tau)\mathrm{d}\tau\\&=\lim_{T\to+\infty}\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau}{2T}\right)(\mathrm{e}^{-\alpha|\tau|}-0)\mathrm{d}\tau\\&=\lim_{T\to+\infty}\frac{1}{T}\int_0^{2T}\left(1-\frac{\tau}{2T}\right)\mathrm{e}^{-\alpha|\tau|}\mathrm{d}\tau\end{aligned}$$

$$= \lim_{T\to+\infty} \frac{1}{T}\left(\frac{1}{\alpha T} - \frac{1-e^{-2\alpha T}}{2\alpha^2 - T^2}\right) = 0$$

关于相关函数 $R_X(\tau)$ 的遍历性定理，可以考虑随机过程 $\{Y_\tau(t), -\infty < t < +\infty\}$，其中 $Y_\tau(t) = X(t+\tau)X(t)$，则 $E[Y_\tau(t)] = R_X(\tau)$，于是由上述均值遍历性定理，可得相关函数的遍历性定理.

定理 2.3　平稳过程 $\{X(t), -\infty < t < +\infty\}$ 的自相关函数具备遍历性的充要条件是

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^{2T}\left(1 - \frac{\tau_1}{2\tau}\right)[B(\tau_1) - R_X^2(\tau)]d\tau_1 = 0$$

式中

$$B(\tau_1) = E[X(t+\tau+\tau_1)X(t+\tau)X(t+\tau_1)X(t)]$$

【例 2.20】　设有 $\{X(t) = \xi\cos(\beta t + \eta), -\infty < t < +\infty\}$，其中 $\xi \sim N(0,1)$，$\eta \sim U(0,2\pi)$，ξ,η 相互独立，试证：$\{X(t), -\infty < t < +\infty\}$ 的均值具有遍历性.

证明

(1) $m_X(t) = E[X(t)] = E[\xi\cos(\beta t + \eta)]$

$$= E(\xi)E[\cos(\beta t + \eta)] = 0$$

(2) $R_X(t, t+\tau) = E[X(t)X(t+\tau)]$

$$= E(\xi^2)E\{\cos(\beta t + \eta)\cos[\beta(t+\tau) + \eta]\}$$

$$= \frac{1}{2}E[\cos(\beta\tau) + \cos(2\beta t + \beta\tau + 2\eta)]$$

$$= \frac{1}{2}\cos(\beta\tau) = R_X(\tau)$$

(3) $E[X^2(t)] = R_X(t,t) = \frac{1}{2} < +\infty$，

由(1)，(2)，(3) 知 $\{X(t), -\infty < t < +\infty\}$ 是平稳过程，又因为

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^{2T}\left(1 - \frac{\tau}{2T}\right)[R_X(\tau) - 0]d\tau$$

$$= \lim_{T\to+\infty} \frac{1}{T}\int_0^{2T}\left(1 - \frac{\tau}{2T}\right)\left[\frac{1}{2}\cos(\beta\tau)\right]d\tau$$

$$= \lim_{T\to+\infty} \frac{1}{2T}\left[\frac{1}{\omega}\left(1 - \frac{\tau}{2T}\right)\sin(\omega\tau) - \frac{1}{2\omega^2 T}\cos(\omega\tau)\right]\Bigg|_0^{2T}$$

$$= \lim_{T\to+\infty} \frac{1}{4\omega^2 T}[1 - \cos(2\omega T)]$$

$$= 0$$

所以 $\{X(t), -\infty < t < +\infty\}$ 的均值具有遍历性.

习　题　2

1. 随机过程$\{X(t),t>0\}$的一维概率密度函数为
$$F_t(x)=P\{X(t)\leqslant x\}=1-\mathrm{e}^{-(x/t)^2},x\geqslant 0$$
此过程是弱平稳的吗？

2. 随机过程$\{X(t),t>0\}$的一维概率密度函数为
$$F_t(x)=P\{X(t)\leqslant x\}=\frac{1}{\sqrt{2\pi t}\sigma}\int_{-\infty}^{x}\mathrm{e}^{-\frac{(u-\mu t)^2}{2\sigma^2 t}}\mathrm{d}u$$
其中$\mu>0,\sigma>0,x\in(-\infty,+\infty)$. 求它的均值函数$m_X(t)$.

3. 令$X(t)=A\sin(\omega t+\Phi)$，其中$A$和$\Phi$是相互独立的非负随机变量，且$\Phi$服从区间$[0,2\pi]$上的均匀分布，$E(A^2)<+\infty$.
(1) 求$\{X(t),t\in(-\infty,+\infty)\}$的均值函数，协方差函数和相关函数.
(2) 随机过程$\{X(t),t\in(-\infty,+\infty)\}$是弱平稳的还是强平稳的？

4. 令$X(t)=A(t)\sin(\omega t+\Phi)$，其中对$\forall t,A(t)$和$\Phi$是相互独立的非负随机变量，$\Phi$服从区间$[0,2\pi]$上的均匀分布. 证明：如果$\{A(t),t\in(-\infty,+\infty)\}$是弱平稳过程，那么随机过程$\{X(t),t\in(-\infty,+\infty)\}$也是弱平稳过程.

5. $\{a_1,a_2,\cdots,a_n\}$是实数序列，$\{\Phi_1,\Phi_2,\cdots,\Phi_n\}$是相互独立的，且服从区间$[0,2\pi]$上均匀分布的随机变量组成的序列. 求随机过程$\{X(t),t\in(-\infty,+\infty)\}$的协方差函数和相关函数，其中$X(t)=\sum_{i=1}^{n}a_i\sin(\omega t+\Phi_i)$.

6. 随机过程$\{X(t)=\sum_{-\infty}^{+\infty}A_n h(t-n),-\infty<t<+\infty\}$，其中$A_n$是相互独立同分布的随机变量，只取值$-1$和$+1$，均值为0，且$h(t)=\begin{cases}1, & 0\leqslant t<\frac{1}{2}\\ 0, & 其他\end{cases}$.
(1) 画出过程$\{X(t),t\in(-\infty,+\infty)\}$的一条可能的样本路径；
(2) 求此过程的协方差函数；
(3) 令$Y(t)=X(t-Z)$，其中随机变量Z服从区间$[0,1]$上的均匀分布，随机过程$\{Y(t),t\in(-\infty,+\infty)\}$是弱平稳的吗？

7. $\{X(t),t\in(-\infty,+\infty)\}$和$\{Y(t),t\in(-\infty,+\infty)\}$是两个独立的弱平稳随机过程，它们的均值函数均为0，且有相同的协方差函数$C_X(\tau)$.
证明：随机过程$\{Z(t),t\in(-\infty,+\infty)\}$是弱平稳的，其中
$$Z(t)=X(t)\cos(\omega t)-Y(t)\sin(\omega t)$$

8. $X(t)=\sin(\Phi t)$，其中Φ在区间$[0,2\pi]$上服从均匀分布，证明：
(1) 离散时间随机过程$\{X(t),t=1,2,\cdots\}$是弱平稳的，但不是强平稳的；

(2) 连续时间随机过程$\{X(t),t\geqslant 0\}$既不是弱平稳也不是强平稳的.

9. $\{X(t),t\in(-\infty,+\infty)\}$和$\{Y(t),t\in(-\infty,+\infty)\}$是两个独立的随机过程，它们的均值函数和协方差函数分别为$m_X(t)$、$m_Y(t)$和$C_X(s,t)$、$C_Y(s,t)$. 令$U(t)=X(t)+Y(t)$, $V(t)=X(t)-Y(t)$, $t\in(-\infty,+\infty)$. 求随机过程$\{U(t),t\in(-\infty,+\infty)\}$和$\{V(t),t\in(-\infty,+\infty)\}$的协方差函数.

10. 随机过程$\{X(t)=A\sin t+B\cos t,t\in\mathbf{R}\}$，其中，$A$, B是均值为0且不相关的随机变量，且$E(A^2)=E(B^2)$. 试讨论$X(t)$的遍历性.

11. 设X_0的概率密度函数是$f(x)=2xI_{(x\geqslant 0)}$，对$n\geqslant 1$, $X_0,X_1,\cdots,X_n,X_{n+1}$是$(1-X_n,1)$上的均匀分布. 试证：$\{X_n,n\geqslant 0\}$是具有均值遍历性的平稳过程.

12. 设平稳过程$\{X(t),-\infty<t<+\infty\}$, $E[X(t)]=0$, $R_X(\tau)=Ae^{-a|\tau|}(1+a|\tau|)$，其中，$A$, a是正常数. 试问$X(t)$的均值是否具有遍历性.

13. 随机过程$\{X(t),t\in\mathbf{R}\}$在每个长度为T的区间$[(n-1)T,nT]$, $n=0,\pm 1,\pm 2,\cdots$上的取值为1或-1，且$P\{X(t)=1\}=P\{X(t)=-1\}=\dfrac{1}{2}$, $(n-1)T<t<nT$，且在不同区间上的取值是独立的，问$\{X(t),t\in\mathbf{R}\}$是否为平稳过程，其均值是否具有遍历性?

第 3 章　Poisson 过程

Poisson 过程是一类较为简单的时间连续、状态离散的随机过程，最早于 1837 年由法国数学家 Poisson 引入，并以他的名字命名. Poisson 过程是一类最重要且应用广泛的计数过程，很多具有独立增量性和平稳增量性的计数过程，只要在同一时刻没有两个或两个以上的事件同时发生，都是 Poisson 过程.

3.1　齐次 Poisson 过程

定义 3.1　（计数过程）$N(t)$ 表示 $(0,t]$ 内已发生事件的总数，如果 $N(t)$ 满足：

(1) $N(t) \geqslant 0$，且取整数值；

(2) $\forall s,t > 0, s < t$，有 $N(s) \leqslant N(t)$ 且 $N(t) - N(s)$ 表示在时间间隔 $(s,t]$ 内事件出现的次数.

则称随机过程 $\{N(t), t \geqslant 0\}$ 为计数过程.

Poisson 过程是计数过程的重要类型之一，其定义如下：

定义 3.2　（齐次 Poisson 过程）如果计数过程 $\{N(t), t \geqslant 0\}$ 满足：

(1) $N(0) = 0$；

(2) $\{N(t), t \geqslant 0\}$ 是独立增量过程；

(3) 增量 $N(s,t) = N(t) - N(s), 0 \leqslant s < t$ 服从参数为 $\lambda(t-s)$ 的 Poisson 分布，即

$$P\{N(s,t) = i\} = \frac{[\lambda(t-s)]^i}{i!} e^{-\lambda(t-s)}; i = 0,1,\cdots \tag{3.1}$$

等价地，设 $[s,t]$ 的区间长度为 $\tau = t - s$，对任意的 $\tau > 0$，有

$$P\{N(s,t) = i\} = \frac{(\lambda\tau)^i}{i!} e^{-\lambda\tau}; i = 0,1,\cdots \tag{3.2}$$

则称计数过程 $\{N(t), t \geqslant 0\}$ 是强度为 $\lambda(\lambda > 0)$ 的齐次 Poisson 过程，一般简称为 Poisson 过程.

由定义 3.2 中的条件(1) 可知，随机事件是从 0 时刻开始计数的，条件(3) 是过程成为 Poisson 过程的直接理由，式(3.2) 表明齐次 Poisson 过程是平稳增量过程，

且 $E[N(t)]=\lambda t \Rightarrow \lambda=\frac{E[N(t)]}{t}$,所以可认为 λ 是单位时间内事件发生的平均次数,故称 λ 为 Poisson 过程的强度或速率,它表明随机事件发生的频繁程度.

【例 3.1】 在对随机服务系统中排队现象的研究中,经常用到 Poisson 过程的模型,例如到达某服务设施的人数,到达电话总机的呼叫次数,都可以用 Poisson 过程来描述. 例如顾客以 $\lambda=4$ 人 /h 的速率到达某商店,该商店上午 9:00 开门. 试求 9:30 时仅到一位顾客,而到 11:30 时总计已到达 5 位顾客的概率.

解 以小时为计时单位,并以 9:00 作为起始时刻,所求事件可表示为 $\{N(0.5)=1,N(2.5)=5\}$. 其概率为

$$
\begin{aligned}
&P\{N(0.5)=1,N(2.5)=5\}\\
&=P\{N(0.5)=1,N(2.5)-N(0.5)=4\}\\
&=\frac{4\times 0.5\mathrm{e}^{-4\times 0.5}}{1!}\times\frac{(4\times 2)^4\mathrm{e}^{-4\times 2}}{4!}=0.0155
\end{aligned}
$$

【例 3.2】 通过十字路口的车流可看作 Poisson 过程,如果 1 min 内没有车通过的概率为 0.2,求 2 min 内有多于 1 辆车通过的概率.

解 设 $N(t)$ 表示$[0,t]$ 内通过的车辆数,$\{N(t),t\geqslant 0\}$ 是 Poisson 过程,则

$$P\{N(t)=k\}=\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}$$

$P\{N(1)=0\}=\mathrm{e}^{-\lambda}=0.2$,所以 $\lambda=\ln 5$,从而有

$$
\begin{aligned}
&P\{N(2)>1\}\\
&=1-P\{N(2)\leqslant 1\}\\
&=1-P\{N(2)=1\}-P\{N(2)=0\}\\
&=1-\frac{(2\lambda)^1}{1!}\mathrm{e}^{-2\lambda}-\frac{(2\lambda)^0}{0!}\mathrm{e}^{-2\lambda}\\
&=0.83
\end{aligned}
$$

为什么实际中有这么多现象可以用 Poisson 过程来描述呢?这是根据小概率事件原理推导出来的. 在第 1 章中已经指出,n 重 Bernouli 试验中,如果每次试验成功的概率 p 很小,而试验的次数 n 很大时,二项分布会逼近 Poisson 分布. 这一想法很自然地可以推广到随机过程的情况. 在很短的时间内事件发生的概率很小,但假如考虑很多个这样很短时间的连接,事件的发生将会有一个大致稳定的概率. 为此,引入 Poisson 过程的第二种定义.

定义 3.3 称一个 $N(0)=0$ 的计数过程$\{N(t),t\geqslant 0\}$ 为强度为 $\lambda(\lambda>0)$ 的齐次 Poisson 过程,如果它满足下面三个条件:

(1)$\{N(t),t\geqslant 0\}$ 是平稳独立增量过程;

(2) 当 $h\to 0$ 时,$P\{N(t,t+h)\geqslant 2\}=o(h)$;

(3)$P\{N(t,t+h)=1\}=\lambda h+o(h)$.

从逼近的观点看，条件(1) 说明试验是独立的，且在每个长度相同的小区间上事件发生的概率是相同的；条件(2) 表明事件是一件一件发生的，在同一瞬间同时发生多个事件的可能性很小；条件(3) 说明事件发生的概率 $p = \lambda h$，而且 p 很小，不发生的概率为 $1 - p = 1 - \lambda h$. 这正好是 Bernouli 试验的模型.

若观测区间是$[0,t]$，将观测区间分成 N 个长度相同的区间，每个区间的长度 $h = \frac{t}{N}$，则当 $N \to +\infty$ 时，$h \to 0$，且由条件(2)，(3) 可知，每个区间可看成一次 Bernouli 试验，而由条件(1) 可知，这 N 次 Bernouli 试验是独立同分布的，从而$[0, t]$ 上事件发生的次数服从二项分布，且当 $N \to +\infty$ 时，二项分布的极限是 Poisson 分布.

定义 3.3 与定义 3.2 相比，更容易应用到实际问题中，作为判定某一现象是否能用 Poisson 过程来描述的依据. 因为定义 3.3 中的条件(1)，(2)，(3) 只需通过基于自然规律的定性推理即可得到验证，而不需要定量调查；而一般很难验证定义 3.2 中的(3)，定义 3.2 常常应用在理论研究中.

定理 3.1 Poisson 过程的两种定义是等价的.

接下来，给出 Poisson 过程两定义等价的严格的数学证明：

证明 先由定义 3.2 推导定义 3.3，要证明齐次 Poisson 过程满足定义3.3 的条件(1)，(2)，(3)，只需证明其满足条件(2) 和(3) 即可. 由定义 3.2 可得：

$$P\{N(t,t+h) \geqslant 2\} = e^{-\lambda h}\sum_{i=2}^{+\infty}\frac{(\lambda h)^i}{i!} = \lambda^2 h^2 e^{-\lambda h}\sum_{i=0}^{+\infty}\frac{(\lambda h)^i}{(i+2)!} \leqslant \lambda^2 h^2 = o(h)$$

$$\begin{aligned} P\{N(t,t+h) = 1\} &= P\{N(\Delta t) - N(0) = 1\} \\ &= e^{-\lambda\Delta t}\frac{\lambda\Delta t}{1!} = \lambda\Delta t\sum_{n=0}^{+\infty}\frac{(-\lambda\Delta t)^n}{n!} \\ &= \lambda\Delta t[1 - \lambda\Delta t + o(\Delta t)] = \lambda\Delta t + o(\Delta t) \end{aligned}$$

其次，若计数过程满足定义 3.3 中的条件(1)，(2)，(3)，下面证明它也满足定义 3.2 中的三个条件.

可以看出，只需验证 $s = 0$ 时式(3.2) 成立即可. 因此，记

$$p_i(t) = P\{N(0,t) = i\} = P\{N(t) = i\};\ i = 0,1,\cdots$$

下面证明

$$p_i(t) = \frac{(\lambda t)^i}{i!}e^{-\lambda t};\ i = 0,1,\cdots \tag{3.3}$$

因为

$$\begin{aligned} p_0(t+h) &= P\{N(t+h) = 0\} \\ &= P\{N(t) = 0, N(t,t+h) = 0\} \\ &= P\{N(t) = 0\}P\{N(t,t+h) = 0\} \quad [\text{条件(1)}] \\ &= p_0(t)p_0(h) \\ &= p_0(t)(1 - \lambda h) + o(h) \quad [\text{条件(2) 和(3)}] \end{aligned}$$

因此

$$\frac{p_0(t+h)-p_0(t)}{h}=-\lambda p_0(t)+\frac{o(h)}{h}$$

令 $h\to 0$ 得

$$p'_0(t)=-\lambda p_0(t)$$

由 $p_0(0)=1$,解得方程:$p_0(t)=\mathrm{e}^{-\lambda t},t\geqslant 0$

所以 $i=0$ 时式(3.3) 成立,同理,$i\geqslant 1$ 时,有

$$\begin{aligned}p_i(t+h)&=P\{N(t+h)=i\}\\&=P\{N(t)=i,N(t,t+h)=0\}+P\{N(t)=i-1,N(t,t+h)=1\}+\\&\quad\sum_{k=2}^{i}P\{N(t)=i-k,N(t,t+h)=k\}\\&=p_i(t)p_0(h)+p_{i-1}(t)p_1(h)+o(h)\\&=p_i(t)(1-\lambda h)+p_{i-1}(t)\lambda h+o(h)\end{aligned}$$

于是

$$\frac{p_i(t+h)-p_i(t)}{h}=-\lambda[p_i(t)-p_{i-1}(t)]+o(h)$$

令 $h\to 0$ 得

$$p'_i(t)=-\lambda[p_i(t)-p_{i-1}(t)];\ i=1,2,\cdots \tag{3.4}$$

由 $p_0(t)=\mathrm{e}^{-\lambda t}$ 开始,利用归纳法即可证得

$$p_i(t)=\frac{(\lambda t)^i}{i!}\mathrm{e}^{-\lambda t};i=0,1,\cdots$$

【例 3.3】 从以往的观察可知,在$[0,t]$内发生交通事故的个数 $N(t)$ 可以描述为 Poisson 过程$\{N(t),t\geqslant 0\}$,平均每 4 h 发生一次事故,即过程的强度 $\lambda=0.25(\mathrm{h}^{-1})$. 求事件"$[0,10]$ 内最多发生 1 次事故,$[10,16]$ 内最少发生 2 次事故,$[16,24]$ 内没有事故发生"发生的概率 p(单位:h).

解 因为$\{N(t),t\geqslant 0\}$为平稳独立增量过程,于是

$$\begin{aligned}p&=P\{N(10)-N(0)\leqslant 1,N(16)-N(10)\geqslant 2,N(24)-N(16)=0\}\\&=P\{N(10)-N(0)\leqslant 1\}P\{N(16)-N(10)\geqslant 2\}P\{N(24)-N(16)=0\}\\&=P\{N(10)\leqslant 1\}P\{N(6)\geqslant 2\}P\{N(8)=0\}.\end{aligned}$$

又

$$\begin{aligned}P\{N(10)\leqslant 1\}&=P\{N(10)=0\}+P\{N(10)=1\}\\&=\mathrm{e}^{-0.25\times 10}+0.25\times 10\times \mathrm{e}^{-0.25\times 10}=0.2873\\P\{N(6)\geqslant 2\}&=1-\mathrm{e}^{-0.25\times 6}-0.25\times 6\times \mathrm{e}^{-0.25\times 6}=0.4422\\P\{N(8)=0\}&=\mathrm{e}^{-0.25\times 8}=0.1353.\end{aligned}$$

因此,所求概率 $p=0.2873\times 0.4422\times 0.1353=0.0172$.

【例 3.4】 设在时间区间$[0,t]$内到商店的顾客数 $N(t)$ 服从强度为 λ 的

Poisson 过程,每个顾客购买货物的概率为 p,不购货物的概率为 $1-p$,且与其到达时间是独立的,也与其他顾客的购买行为是独立的.令 $Y(t)$ 为 $[0,t]$ 内购买货物的顾客数,证明:$\{Y(t),t\geqslant 0\}$ 是强度为 λp 的 Poisson 过程.

证明　(1) 显然 $Y(0)=0$.

(2) 令 X_i 表示第 i 个顾客是否购买货物的情况,则 X_i 相互独立且服从共同的参数为 p 的两点分布,且 $Y(t)=\sum_{i=1}^{N(t)}X_i$.

$\forall\, 0<t_1<t_2\cdots<t_n$,$Y(t)$ 的一组增量为

$$Y(t_2)-Y(t_1)=\sum_{i=N(t_1)+1}^{N(t_2)}X_i$$

$$\vdots \qquad\qquad \vdots$$

$$Y(t_n)-Y(t_{n-1})=\sum_{i=N(t_{n-1})+1}^{N(t_n)}X_i$$

由 X_i 相互独立和 Poisson 过程 $N(t)$ 的独立增量性可知,增量

$$Y(t_2)-Y(t_1),Y(t_3)-Y(t_2),\cdots,Y(t_n)-Y(t_{n-1})$$

是相互独立的.

(3) $\forall\, t>s>0$

$$\begin{aligned}
P\{Y(t)-Y(s)=k\} &= \sum_{i=k}^{+\infty}P\{N(t)-N(s)=i\}P\{Y(t)-Y(s)=k\mid N(t)-N(s)=i\}\\
&=\sum_{i=k}^{+\infty}\frac{[\lambda(t-s)]^i e^{-\lambda(t-s)}}{i!}C_i^k p^k(1-p)^{i-k}\\
&=\sum_{i=k}^{+\infty}\frac{[\lambda(t-s)]^i e^{-\lambda(t-s)}}{i!}\frac{i!}{k!(i-k)!}p^k(1-p)^{i-k}\\
&=\sum_{i=k}^{+\infty}\frac{[\lambda(t-s)]^k e^{-\lambda(t-s)}p^k}{k!}\frac{[\lambda(t-s)(1-p)]^{i-k}}{(i-k)!}\\
&=\frac{[\lambda(t-s)p]^k e^{-\lambda(t-s)}}{k!}e^{\lambda(t-s)(1-p)}\\
&=\frac{[\lambda(t-s)p]^k e^{-\lambda(t-s)p}}{k!}
\end{aligned}$$

即 $Y(t)-Y(s)$ 服从参数为 $\lambda(t-s)p$ 的 Poisson 分布.

由(1)(2)(3) 可知,$\{Y(t),t\geqslant 0\}$ 为强度为 λp 的 Poisson 过程.

【例 3.5】　设 $\{N(t),t\geqslant 0\}$ 是强度为 λ 的齐次 Poisson 过程,证明:在已知 $N(t)=n$ 的条件下,$[0,s]$ 内事件发生的次数 $N(s)$ 服从二项分布 $B\left(n,\frac{s}{t}\right)$,$s<t$.

证明　因为 $\{N(t),t\geqslant 0\}$ 为平稳增量过程,所以

$$P\{N(s)=k \mid N(t)=n\}=\frac{P\{N(s)=k,N(t)=n\}}{P\{N(t)=n\}}$$
$$=\frac{P\{N(s)=k,N(s,t)=n-k\}}{P\{N(t)=n\}}$$
$$=\frac{P\{N(s)=k\}P\{N(s,t)=n-k\}}{P\{N(t)=n\}}$$
$$=\frac{\frac{(\lambda s)^k}{k!}e^{-\lambda s}\frac{[\lambda(t-s)]^{n-k}}{(n-k)!}e^{-\lambda(t-s)}}{\frac{(\lambda t)^n}{n!}e^{-\lambda t}}$$
$$=\binom{n}{k}\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{n-k};\quad k=0,1,\cdots,n$$

得证.

下面的例子用到了双曲正弦和双曲余弦函数：

$$\sinh x=\frac{e^x-e^{-x}}{2},\quad \cosh x=\frac{e^x+e^{-x}}{2},\quad x\in(-\infty,+\infty)$$

【例 3.6】 (随机电报信号)若一个随机信号 $X(t)$ 满足

$$X(t)=Y(-1)^{N(t)},\quad t\geqslant 0$$

其中$\{N(t),t\geqslant 0\}$是强度为$\lambda(\lambda>0)$的齐次 Poisson 过程,Y为二元随机变量,与$N(t)$独立且$P\{Y=1\}=P\{Y=-1\}=\frac{1}{2}$,则称这种随机信号为随机电报信号.

随机电报信号是产生复杂信号结构的基本单元,明显地,$X(t)=1$或$X(t)=-1$,$X(0)=Y$. 图 3.1 所示为在$Y=1,T_n=t_n;n=1,2,\cdots$的条件下,随机过程$\{X(t),t\geqslant 0\}$的一条样本轨道$x=x(t)$.

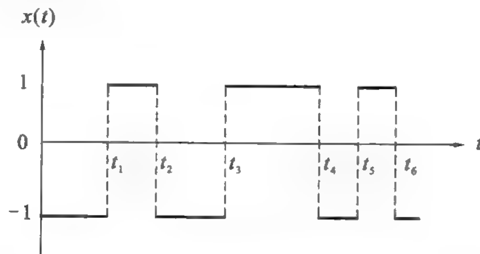


图 3.1 随机电报信号的样本轨道

下证$\{X(t),t\geqslant 0\}$为宽平稳过程：

① 由$|X(t)|^2=Y^2=1<+\infty$知,$\forall t\geqslant 0$,$\{X(t),t\geqslant 0\}$为二阶过程.

② 记$I(t)=(-1)^{N(t)}$,则均值函数

$$m(t)=E[X(t)]=E(Y)E[I(t)]$$

又$E(Y)=0$,因此$m(t)\equiv 0$.

③ $\forall s,\ t>0$

$$\begin{aligned}C_X(s,t) &= \mathrm{Cov}[X(s),X(t)]\\ &= E[X(s)X(t)] = E[YI(s)YI(t)]\\ &= E[Y^2I(s)I(t)] = E(Y^2)E[I(s)I(t)]\end{aligned}$$

而其中 $E(Y^2)=1$，所以 $C_X(s,t)=E[I(s)I(t)]$，要求 $C_X(s,t)$ 的值，需先确定 $(I(s),I(t))$ 的联合分布和 $I(t)$ 的概率分布.

Poisson 事件发生时，$I(t)$ 完成一次从 $I(t)=-1$ 到 $I(t)=1$，或从 $I(t)=1$ 到 $I(t)=-1$ 的转变，即当 $N(t)$ 跳跃时，

$$P\{I(t)=1\}=P\{[0,t]\text{内有偶数次跳跃}\}=\mathrm{e}^{-\lambda t}\sum_{i=0}^{+\infty}\frac{(\lambda t)^{2i}}{(2i)!}=\mathrm{e}^{-\lambda t}\cosh(\lambda t)$$

同理，

$$P\{I(t)=-1\}=P\{[0,t]\text{内有奇数次跳跃}\}=\mathrm{e}^{-\lambda t}\sum_{i=0}^{+\infty}\frac{(\lambda t)^{2i+1}}{(2i+1)!}=\mathrm{e}^{-\lambda t}\sinh(\lambda t)$$

因此，

$$\begin{aligned}E[I(t)] &= 1\cdot P\{I(t)=1\}+(-1)\cdot P\{I(t)=-1\}\\ &= \mathrm{e}^{-\lambda t}[\cosh(\lambda t)-\sinh(\lambda t)]=\mathrm{e}^{-2\lambda t}\end{aligned}$$

利用条件概率和 $\{N(t),t\geqslant 0\}$ 的平稳增量性可知，若 $s<t$，则

$$\begin{aligned}p_{1,1} &= P\{I(s)=1,I(t)=1\}\\ &= P\{I(s)=1\}P\{I(t)=1\mid I(s)=1\}\\ &= \mathrm{e}^{-\lambda s}\cosh(\lambda s)P\{(s,t]\text{内有偶数次跳跃}\}\\ &= \mathrm{e}^{-\lambda s}\cosh(\lambda s)\mathrm{e}^{-\lambda(t-s)}\cosh[\lambda(t-s)]\\ &= \mathrm{e}^{-\lambda t}\cosh(\lambda s)\cosh[\lambda(t-s)]\end{aligned}$$

类似地，

$$\begin{aligned}p_{1,-1} &= P\{I(s)=1,I(t)=-1\}=\mathrm{e}^{-\lambda t}\cosh(\lambda s)\sinh[\lambda(t-s)]\\ p_{-1,1} &= P\{I(s)=-1,I(t)=1\}=\mathrm{e}^{-\lambda t}\sinh(\lambda s)\sinh[\lambda(t-s)]\\ p_{-1,-1} &= P\{I(s)=-1,I(t)=-1\}=\mathrm{e}^{-\lambda t}\sinh(\lambda s)\cosh[\lambda(t-s)]\end{aligned}$$

此时，

$$E[I(s)I(t)]=p_{1,1}+p_{-1,-1}+p_{1,-1}+p_{-1,1}$$

所以，

$$C_X(s,t)=\mathrm{e}^{-2\lambda(t-s)},\quad s<t$$

由于 s 与 t 的顺序可变，故 $C_X(s,t)=\mathrm{e}^{-2\lambda|t-s|}$.

即证 $\{X(t),t\geqslant 0\}$ 为弱平稳过程.

3.2 Poisson 过程的可加性和可分解性

假设到达某服务系统的顾客有 n 个不同来源，比如：到达某银行的顾客来自 n 个不同的小镇，共有 n 种类型的汽车到达某汽车保养维修店，则每个小镇顾客或每种类型汽车的达到都是计数过程. 令 $\{N_i(t), t \geqslant 0\}, i = 1,2,\cdots,n$，表示相应的计数过程，则 $[0,t]$ 内达到某服务设施总的顾客数是 $N(t) = N_1(t) + N_2(t) + \cdots + N_n(t), t \geqslant 0$. 如果 $\{N_i(t), t \geqslant 0\}$ 是强度为 λ_i 的齐次 Poisson 过程，且相互独立，那么计数过程 $\{N(t), t \geqslant 0\}$ 是否仍然为 Poisson 过程呢？

定理 3.2 （Poisson 过程的可加性）如果 $\{N_i(t), t \geqslant 0\}, i = 1,2,\cdots,n$ 是相互独立的、强度为 λ_i 的齐次 Poisson 过程，则

$$N(t) = N_1(t) + N_2(t) + \cdots + N_n(t), t \geqslant 0$$

是强度为 $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ 的齐次 Poisson 过程.

证明

(1) $N(0) = N_1(0) + N_2(0) + \cdots + N_n(0) = 0$

(2) $\forall\, 0 < t_1 < t_2 < \cdots < t_{m-1} < t_m$，有

$$N(t_2) - N(t_1) = \sum_{i=1}^{n} [N_i(t_2) - N_i(t_1)]$$

$$\vdots \qquad\qquad \vdots$$

$$N(t_m) - N(t_{m-1}) = \sum_{i=1}^{n} [N_i(t_m) - N_i(t_{m-1})]$$

由 $N_i(t), i = 1,2,\cdots,n$ 为 Poisson 过程可知，增量 $N_i(t_2) - N_i(t_1), N_i(t_3) - N_i(t_2), \cdots, N_i(t_m) - N_i(t_{m-1})$ 是相互独立的，且 $N_i(t), i = 1,2,\cdots,n$ 是独立的 Poisson 过程，从而增量 $N(t_2) - N(t_1), N(t_3) - N(t_2), \cdots, N(t_m) - N(t_{m-1})$ 也是相互独立的.

$\forall\, t_1, t_2 > 0$，有

$$N(t_1 + h) - N(t_1) = \sum_{i=1}^{n} [N_i(t_1 + h) - N_i(t_1)]$$

$$N(t_2 + h) - N(t_2) = \sum_{i=1}^{n} [N_i(t_2 + h) - N_i(t_2)]$$

又

$$N_i(t_1 + h) - N_i(t_1) \stackrel{d}{=} N_i(t_2 + h) - N_i(t_2), i = 1,2,\cdots,n$$

所以

$$N(t_1 + h) - N(t_1) \stackrel{d}{=} N(t_2 + h) - N(t_2)$$

其中“$\stackrel{d}{=}$”表示同分布，即 $N(t)$ 是平稳增量过程.

(3) $N_i(t)$ 的特征函数为

$$f_{N_i}(u) = E[e^{iuN_i(t)}] = \exp[\lambda_i t(e^{iu} - 1)]$$

所以 $N(t)$ 的特征函数为

$$f_N(u) = \prod_{i=1}^{n} f_{N_i}(u) = \exp[(\sum_{i=1}^{n}\lambda_i)t(e^{iu} - 1)]$$

即 $N(t)$ 服从参数为 $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ 的 Poisson 分布.

由(1)(2)(3)可知，$N(t)$ 是强度为 $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$ 的齐次 Poisson 过程.

设旅客按强度为 λ 的 Poisson 过程到达长途汽车站，每次到达的旅客乘 A 线的概率是 p，乘 B 线的概率是 $q = 1 - p$，且与其到达时间相互独立，不同的旅客到达行为也是独立的. 如果 $N_1(t)$ 表示 $[0,t]$ 内乘 A 线旅客到达的人数，$N_2(t)$ 表示 $[0,t]$ 内乘 B 线旅客到达的人数，则由例 3.4 知 $\{N_1(t), t \geq 0\}$ 是强度为 λp 的 Poisson 过程. 类似地，可证明 $\{N_2(t), t \geq 0\}$ 是强度为 $\lambda(1-p)$ 的 Poisson 过程，且 $N_1(t)$ 和 $N_2(t)$ 相互独立. $N_1(t)$ 和 $N_2(t)$ 被称为 $N(t)$ 的分流.

如果考虑乘 A 线的旅客以概率 p_i 前往 A_i 线，则可进一步把 Poisson 过程 $\{N_1(t), t \geq 0\}$ 再进行分流，于是可得到下面的定理：

定理 3.3 （Poisson 过程的可分解性）$N(t)$ 是强度为 λ 的齐次 Poisson 过程，可分解为 m 个相互独立的，强度分别为 $\lambda p_i, i = 1,2,\cdots,m$ 的 Poisson 过程，其中 $0 < p_i < 1, \sum_{i=1}^{m} p_i = 1$.

证明略.

【例 3.7】 汽车按 Poisson 过程驶向立体交叉桥，如果由东面平均每分钟驶入 6 辆汽车，由南面平均每分钟驶入 6.5 辆汽车，由西面平均每分钟驶入 9 辆汽车，由北面平均每分钟驶入 8.5 辆汽车；在该桥上，每辆车向左或向右转向行驶的概率都是0.30，直行的概率是 0.35，掉头行驶的概率是 0.05. 计算各个方向上，离开立交桥汽车流的强度.

解 用 $N_1(t)$ 表示 $[0,t]$ 内由东面驶入的汽车辆数，由题意知 $\{N_1(t), t \geq 0\}$ 是强度为 $\lambda_1 = 6$(辆 /min) 的 Poisson 过程. 根据 Poisson 过程的可分解性，由东面驶入的车流分解到东、南、西、北四个方向的车流分别是强度为 $0.05\lambda_1, 0.30\lambda_1, 0.35\lambda_1, 0.30\lambda_1$ 的 Poisson 过程.

可类似地列出其他方向车流的分解情况，列入表 3.1 中.

表 3.1 车流分解情况

方向	向东分解	向南分解	向西分解	向北分解
东面驶入 $\lambda_1 = 6.0$	$0.05\lambda_1$	$0.30\lambda_1$	$0.35\lambda_1$	$0.30\lambda_1$

续表 3.1

方向	向东分解	向南分解	向西分解	向北分解
南面驶入 $\lambda_2 = 6.5$	$0.30\lambda_2$	$0.05\lambda_2$	$0.30\lambda_2$	$0.35\lambda_2$
西面驶入 $\lambda_3 = 9.0$	$0.35\lambda_3$	$0.30\lambda_3$	$0.05\lambda_3$	$0.30\lambda_3$
北面驶入 $\lambda_4 = 8.5$	$0.30\lambda_4$	$0.35\lambda_4$	$0.30\lambda_4$	$0.05\lambda_4$
驶出强度(辆/min)	$\lambda_E = 7.95$	$\lambda_S = 7.80$	$\lambda_W = 7.05$	$\lambda_N = 7.20$

根据Poisson过程的可加性,得出表中最后一行,其中$\lambda_E = 7.95$是向东驶出立交桥的汽车流强度,是λ_E所在列的各过程的强度之和;$\lambda_S = 7.80$是向南驶出立交桥的汽车流强度,是λ_S所在列的各过程的强度之和;$\lambda_W = 7.05$是向西驶出立交桥的汽车流强度,是λ_W所在列的各过程的强度之和;$\lambda_N = 7.20$是向北驶出立交桥的汽车流强度,是λ_N所在列的各过程的强度之和.各个方向上,离开立交桥的汽车流都是Poisson过程.

3.3 Poisson过程与指数分布

设$\{N(t), t \geqslant 0\}$是Poisson过程,令$T_0 = 0$,T_n表示第n个事件发生的时刻或第n个事件发生的等待时间($n \geqslant 1$),$X_n = T_n - T_{n-1}$($n \geqslant 1$)表示第$n-1$个事件与第n个事件发生的时间间隔.

当$\forall t \geqslant 0, n \geqslant 0$时,下列事件等价:

$$\{N(t) \geqslant n\} = \{T_n \leqslant t\}$$

$$\{N(t) = n\} = \{T_n \leqslant t < T_{n+1}\} = \{T_n \leqslant t\} - \{T_{n+1} \leqslant t\}$$

因此,当$t < 0$时,$F_{T_n}(t) = 0$;

当$t \geqslant 0$时,有

$$F_{T_n}(t) = P\{T_n \leqslant t\} = P\{N(t) \geqslant n\} = \mathrm{e}^{-\lambda t}\sum_{k=n}^{+\infty}\frac{(\lambda t)^k}{k!}, n = 1, 2, \cdots$$

两端同时对t求导可得T_n的密度函数

$$f_{T_n}(t) = \lambda \mathrm{e}^{-\lambda t}\sum_{i=n}^{+\infty}\frac{(\lambda t)^{i-1}}{(i-1)!} - \lambda \mathrm{e}^{-\lambda t}\sum_{i=n}^{+\infty}\frac{(\lambda t)^i}{i!}$$

$$= \frac{(\lambda t)^{n-1}}{(n-1)!}\lambda \mathrm{e}^{-\lambda t}$$

即$T_n \sim \Gamma(n, \lambda)$.

特别地,当$n = 1$时,有

$$P\{X_1 \leqslant t\} = P\{T_1 \leqslant t\} = 1 - \mathrm{e}^{-\lambda t}, t \geqslant 0$$

即$X_1 \sim \mathrm{e}(\lambda)$.

问 $X_2,X_3,\cdots,X_n$ 是否还是服从指数分布?现给出一个重要的定理:

定理 3.4 设$\{N(t),t\geqslant 0\}$是 Poisson 过程,$T_0=0$,T_n 表示第 n 个事件发生的时刻($n\geqslant 1$),$X_n=T_n-T_{n-1}(n\geqslant 1)$.则 $X_1,X_2,\cdots$ 相互独立同分布,其共同分布是参数为 λ 的指数分布.

证明 由上面的证明可知 $X_1\sim e(\lambda)$,对 $\forall s<t$,有

$$\begin{aligned}P\{X_2>t\mid X_1=s\}&=P\{(s,s+t]\text{上没有事件发生}\mid X_1=s\}\\&=P\{N(s+t)-N(s)=0\mid N(s)=1\}\\&=P\{N(s+t)-N(s)=0\}\\&=P\{N(t)-N(0)=0\}\\&=e^{-\lambda t}\end{aligned}$$

所以 $P\{X_2>t\}=e^{-\lambda t}$,更进一步有 X_1 与 X_2 相互独立.

对于 $\forall n\geqslant 1$ 和 $t,s_1,s_2,\cdots,s_{n-1}\geqslant 0$,有

$$\begin{aligned}&P\{X_n>t\mid X_1=s_1,X_2=s_2,\cdots,X_{n-1}=s_{n-1}\}\\&=P\{(s_1+s_2+\cdots+s_{n-1},s_1+s_2+\cdots+s_{n-1}+t]\text{内没有事件发生}\\&\qquad\mid X_1=s_1,X_2=s_2,\cdots,X_{n-1}=s_{n-1}\}\\&=P\{N(\sum_{i=1}^{n-1}s_i+t)-N(\sum_{i=1}^{n-1}s_i)=0\mid N(s_1)=1,\\&\qquad N(s_1+s_2)-N(s_1)=1,\cdots,N(\sum_{i=1}^{n-1}s_i)-N(\sum_{i=1}^{n-2}s_i)=1\}\\&=P\{N(\sum_{i=1}^{n-1}s_i+t)-N(\sum_{i=1}^{n-1}s_i)=0\}\\&=P\{N(t)-N(0)=0\}\\&=e^{-\lambda t}\end{aligned}$$

从而有 $P\{X_n>t\}=e^{-\lambda t}$,所以 $X_1,X_2,\cdots$ 都服从参数为 λ 的指数分布,且相互独立.

Poisson 过程具有平稳独立增量性,其概率意义是指从任何时刻起,过程独立于之前发生的一切,且与原过程有完全相同的分布,而具有无记忆性的连续型分布只有指数分布,所以时间间隔服从指数分布是预料之中的.

定理 3.4 的逆命题也是成立的,这又给定义 Poisson 过程带来了另一种方法.

定理 3.5 若计数过程$\{N(t),t\geqslant 0\}$每次事件发生的时间间隔 $X_1,X_2,\cdots$ 相互独立,且服从同一参数为 λ 的指数分布,则$\{N(t),t\geqslant 0\}$是强度为 λ 的 Poisson 过程.

证明略.

定理 3.5 提供了对 Poisson 过程进行计算机模拟的便捷途径.大家对均匀分布

$U[0,1]$ 的随机抽样方法是熟知的，将其做变换后就可以模拟参数为 λ 的指数随机变量，而独立同分布的指数随机变量的和 $\sum_{i=1}^{n} X_n$ 就是 T_n，也就是 Poisson 过程第 n 个事件的发生时刻，这样就可以模拟参数为 λ 的 Poisson 过程. λ 越大说明事件发生的平均间隔时间 $\frac{1}{\lambda}$ 越短，事件的发生就越频繁，强度也就越大.

【例 3.8】 放射性物质在衰减过程中，平均每分钟放射出 4 个 γ 光子，用 $N(t)$ 表示在观测时间区间 $(0,t]$ 内放射出 γ 光子的数目，且 $\{N(t),t\geqslant 0\}$ 是 Poisson 过程. 设计数器对检测到的 γ 光子只是每隔一个记录一次，令 T 是两个相继被记录的光子之间的时间间隔(以 min 为单位)，求 T 的概率密度函数.

解 由题意，$E[N(1)]=4=\lambda$，故 $\{N(t),t\geqslant 0\}$ 是参数为 $\lambda=4$ 的 Poisson 过程.

设 $Z_k,k=1,2,\cdots$ 表示第 $k-1$ 个与第 k 个被记录的光子之间的时间间隔，且从放射出的第 2 个光子开始记录，X_k 表示放射出的光子之间的时间间隔，显然 $Z_k=X_{2k-1}+X_{2k}$. 由定理 3.4 知，$X_k,k=1,2,\cdots$ 独立且同指数分布，于是 $Z_k,k=1,2,\cdots$ 也是独立同分布的. 所以只要求出 Z_1 的分布，即为 T 的分布. 注意到 $\{Z_1>t\}=\{$在$[0,t)$ 内至多到达 1 个光子$\}$，故

$$P\{Z_1>t\}=P\{X_1+X_2>t\}=P\{N(t)\leqslant 1\}=P\{N(t)=0\}+P\{N(t)=1\}$$
$$=e^{-4t}+4te^{-4t}=(1+4t)e^{-4t},t\geqslant 0$$

所以 T 的分布函数为

$$F_T(t)=P\{Z_1\leqslant t\}=\begin{cases}1-(1+4t)e^{-4t}, & t\geqslant 0\\ 0, & t<0\end{cases}$$

概率密度函数为

$$f_T(t)=F'_T(t)=\begin{cases}16te^{-4t}, & t\geqslant 0\\ 0, & t<0\end{cases}$$

3.4 Poisson 过程与均匀分布

例 3.5 表明在已知 $N(t)=1$ 的条件下，第一次事件发生的时刻 T_1 在$[0,t]$上是均匀分布的，因为，对于 $s<t$，有

$$P\{T_1\leqslant s|N(t)=1\}=P\{N(s)=1\mid N(t)=1\}=\frac{s}{t}$$

这只是齐次 Poisson 过程与均匀分布之间关系的一个特例，为了推广这一结论，先引入顺序统计量的概念及其相关结论.

设$(X_1,X_2,\cdots,X_n)$为来自 X 的随机样本，即 X_i 相互独立，且与随机变量 X 具有相同的分布. 其相应的顺序统计量记为

$$(X_1^*, X_2^*, \cdots, X_n^*), \quad 0 \leqslant X_1^* \leqslant X_2^* \leqslant \cdots \leqslant X_n^*$$

若 X 在 $[0,x]$ 上均匀分布，则 $(X_1^*, X_2^*, \cdots, X_n^*)$ 的联合概率密度为

$$f^*(x_1^*, x_2^*, \cdots, x_n^*) = \begin{cases} \dfrac{n!}{x^n}, & 0 \leqslant x_1^* \leqslant x_2^* \leqslant \cdots \leqslant x_n^* \\ 0, & \text{其他} \end{cases}$$

而未排序的样本 $(X_1, X_2, \cdots, X_n)$ 的联合概率密度为

$$f(x_1, x_2, \cdots, x_n) = \begin{cases} \dfrac{1}{x^n}, & 0 \leqslant x_i < x \\ 0, & \text{其他} \end{cases}$$

定理 3.6 设 $\{N(t), t \geqslant 0\}$ 是强度为 λ 的齐次 Poisson 过程，$T_0 = 0$，T_n 表示第 n 个事件发生的时刻 $(n \geqslant 1)$. 在已知 $N(t) = n, n = 1, 2, \cdots$ 的条件下，$(T_1, T_2, \cdots, T_n)$ 与 $[0,t]$ 上服从均匀分布的 n 个随机变量的顺序统计量的联合分布密度函数相同，即

$$f(t_1, t_2, \cdots, t_n \mid N(t) = n) = \begin{cases} \dfrac{n!}{t^n}, & 0 \leqslant t_1 < t_2 < \cdots < t_n \leqslant t \\ 0, & \text{其他} \end{cases}$$

证明 对 $\forall\, 0 < t_1 < t_2 < \cdots < t_n < t_{n+1} = t$，取 $h_0 = h_{n+1} = 0$ 及充分小的 h_i，使得 $t_i + h_i < t_{i+1}$，$1 \leqslant i \leqslant n$，则有

$$\begin{aligned} & P\{t_i < T_i \leqslant t_i + h_i, 1 \leqslant i \leqslant n \mid N(t) = n\} \\ &= \frac{P\{(t_i, t_i + h_i] \text{ 中有一事件发生}, 1 \leqslant i \leqslant n, [0,t] \text{ 的别处无事件发生}\}}{P\{N(t) = n\}} \\ &= \frac{P\{N(t_i + h_i) - N(t_i) = 1, N(t_{i+1}) - N(t_i + h_i) = 0, 1 \leqslant i \leqslant n, N(t_1) = 0\}}{P\{N(t) = n\}} \\ &= \frac{(\lambda h_1)\mathrm{e}^{-\lambda h_1}(\lambda h_2)\mathrm{e}^{-\lambda h_2} \cdots (\lambda h_n)\mathrm{e}^{-\lambda h_n}\mathrm{e}^{-\lambda(t - h_1 - h_2 - \cdots - h_n)}}{\dfrac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}} = \frac{n!}{t^n} h_1 h_2 \cdots h_n \end{aligned}$$

由此得，在已知 $N(t) = n$ 的条件下，当 $0 \leqslant t_1 < t_2 < \cdots < t_n \leqslant t$ 时，$(T_1, T_2, \cdots, T_n)$ 的联合分布密度函数为

$$\begin{aligned} f(t_1, t_2, \cdots, t_n \mid N(t) = n) &= \lim_{\substack{h_i \to 0 \\ 1 \leqslant i \leqslant n}} \frac{P\{t_i < T_i \leqslant t_i + h_i, 1 \leqslant i \leqslant n \mid N(t) = n\}}{h_1 h_2 \cdots h_n} \\ &= \frac{n!}{t^n} \end{aligned}$$

所以

$$f(t_1, t_2, \cdots, t_n \mid N(t) = n) = \begin{cases} \dfrac{n!}{t^n}, & 0 \leqslant t_1 < t_2 < \cdots < t_n \leqslant t \\ 0, & \text{其他} \end{cases}$$

定理 3.6 说明在 $N(t) = n$ 的条件下，事件相继发生的时间 $T_1, T_2, \cdots, T_n$ 的条

件分布函数,与 n 个在$[0,t]$上相互独立且同服从均匀分布的随机变量的顺序统计量的分布密度函数一样.所以直观上,在已知在$[0,t]$内发生了 n 次事件的条件下,各次事件发生的时刻 $T_1,T_2,\cdots,T_n$(不排序)可看作是相互独立的、服从$[0,t]$上均匀分布的随机变量.

【例3.9】 乘客按照强度为λ的Poisson过程来到某火车站,火车在时刻t离开车站,计算在$[0,t]$内到达车站的顾客等待时间总和的期望值.

解 设第 i 个乘客到达火车站的时刻为 T_i,他的等待时间为 $t-T_i$,在$[0,t]$内总共来了 $N(t)$ 位乘客,则$[0,t]$内到达车站的顾客等待时间总和为

$$S(t)=\sum_{i=1}^{N(t)}(t-T_i)$$

因为

$$\begin{aligned}E[S(t)\mid N(t)=n]&=E\Big[\sum_{i=1}^{N(t)}(t-T_i)\,\Big|\,N(t)=n\Big]\\&=E\Big[\sum_{i=1}^{n}(t-T_i)\,\Big|\,N(t)=n\Big]=nt-E\Big[\sum_{i=1}^{n}T_i\,\Big|\,N(t)=n\Big]\\&=nt-E\Big[\sum_{i=1}^{n}U_{(i)}\Big]\\&=nt-E\Big[\sum_{i=1}^{n}U_i\Big]\\&=nt-\frac{nt}{2}=\frac{nt}{2}\end{aligned}$$

其中 $U_1,U_2,\cdots,U_n$ 为$[0,t]$上独立且同均匀分布的随机变量,$U_{(1)}\leqslant U_{(2)}\leqslant\cdots\leqslant U_{(n)}$ 为其顺序统计量.故

$$E[S(t)]=E\{E[S(t)\mid N(t)]\}=\frac{t}{2}E[N(t)]=\frac{\lambda}{2}t^2$$

【例3.10】 设一部仪器在$[0,t]$内受冲击的次数$\{N(t),t\geqslant 0\}$是参数为λ的齐次Poisson过程.第 i 次冲击造成的损失为 D_i,假定 $D_i,i\geqslant 1$ 独立同分布且与$\{N(t),t\geqslant 0\}$独立,冲击引起的损失随时间按负指数衰减,即 $t=0$ 的衰减为 D,经时刻 t 损失为 $De^{-\alpha t}$($\alpha>0$ 为常数).设损失是可加的,那么在 t 时刻的总损失为 $D(t)=\sum_{i=1}^{N(t)}D_i e^{-\alpha(t-T_i)}$,其中 T_i 为第 i 次冲击到达的时刻,试求 $E[D(t)]$.

解 由于

$$\begin{aligned}E[D(t)\mid N(t)=n]&=E\Big[\sum_{i=1}^{N(t)}D_i e^{-\alpha(t-T_i)}\,\Big|\,N(t)=n\Big]\\&=E\Big[\sum_{i=1}^{n}D_i e^{-\alpha(t-T_i)}\,\Big|\,N(t)=n\Big]\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} E[D_i e^{-\alpha(t-T_i)} \mid N(t)=n] \\
&= \sum_{i=1}^{n} E[D_i \mid N(t)=n] E[e^{-\alpha(t-T_i)} \mid N(t)=n] \\
&= E(D) \cdot e^{-\alpha t} \sum_{i=1}^{n} E[e^{\alpha T_i} \mid N(t)=n]
\end{aligned}
$$

记$U_1, U_2, \cdots, U_n$为$[0,t]$上独立且同均匀分布的随机变量，则有

$$
\begin{aligned}
\sum_{i=1}^{n} E[e^{\alpha T_i} \mid N(t)=n] &= E[\sum_{i=1}^{n} e^{\alpha U_{(i)}}] = E[\sum_{k=1}^{n} e^{\alpha U_i}] \\
&= n\int_0^t e^{\alpha x} \frac{1}{t} dx \\
&= \frac{n}{\alpha t}(e^{\alpha t} - 1)
\end{aligned}
$$

所以有

$$
E[X(t) \mid N(t)=n] = \frac{n}{\alpha t}(1-e^{-\alpha t})E(D)
$$

即有

$$
E[X(t) \mid N(t)] = \frac{N(t)}{\alpha t}(1-e^{-\alpha t})E(D)
$$

故

$$
E[X(t)] = E\{E[X(t) \mid N(t)]\} = \frac{\lambda E(D)}{\alpha}(1-e^{-\alpha t})
$$

【例3.11】 假设参数为λ的Poisson过程的各个事件被分为1类或2类，具体地说，如果某一事件在s时刻发生以概率$P(s)$归为1类，以概率$1-P(s)$归为2类，且与其他事件归为哪一类相互独立. 若$N_i(t)$表示到时刻t，i类事件发生的件数，$i=1,2$，则$N_1(t)$与$N_2(t)$是独立的，均值分别为λtp，$\lambda t(1-p)$的泊松随机变量，其中

$$
p = \frac{1}{t}\int_0^t P(s) ds
$$

证明　由全概率公式得

$$
\begin{aligned}
&P\{N_1(t)=n, N_2(t)=m\} \\
&= \sum_{k=0}^{+\infty} P\{N_1(t)=n, N_2(t)=m \mid N(t)=k\} P\{N(t)=k\} \\
&= P\{N_1(t)=n, N_2(t)=m \mid N(t)=m+n\} P\{N(t)=m+n\}
\end{aligned}
$$

考虑在$[0,t]$上发生的任一事件，如果它在s时刻发生，则它是1类事件的概率为$P(s)$. 而由定理3.6知，此事件发生的时刻在$(0,t)$上均匀分布，所以它是1类事件的概率为

$$
p = \frac{1}{t}\int_0^t P(s) ds
$$

且与其他事件是哪一类相互独立.

因此,$P\{N_1(t)=n,N_2(t)=m \mid N(t)=m+n\}$ 可看成 $m+n$ 次 Bernouli 试验中 n 次成功,m 次失败的概率,而每次试验成功的概率是 p,即

$$P\{N_1(t)=n,N_2(t)=m \mid N(t)=m+n\}=\binom{m+n}{n}p^n(1-p)^m$$

从而

$$\begin{aligned}
&P\{N_1(t)=n,N_2(t)=m\}\\
&=P\{N_1(t)=n,N_2(t)=m \mid N(t)=m+n\}P\{N(t)=m+n\}\\
&=\binom{m+n}{n}p^n(1-p)^m\frac{(\lambda t)^{m+n}}{(m+n)!}e^{-\lambda t}\\
&=\frac{(\lambda tp)^n}{n!}e^{-\lambda tp}\frac{[\lambda t(1-p)]^m}{m!}e^{-\lambda t(1-p)}
\end{aligned}$$

所以

$$P\{N_1(t)=n\}=\frac{(\lambda tp)^n}{n!}e^{-\lambda tp}$$

$$P\{N_2(t)=m\}=\frac{[\lambda t(1-p)]^m}{m!}e^{-\lambda t(1-p)}$$

且 $N_1(t)$ 与 $N_2(t)$ 相互独立.

【例 3.12】 ($M/G/\infty$ 排队系统)$M/G/\infty$ 表示一个随机服务系统,其中 M 表示顾客到达服务台的过程是参数为 λ 的 Poisson 过程;G 表示每个顾客到达后的服务时间 Y 是独立同分布的随机变量,其分布函数为 $G(t)$;∞ 表示服务人员的人数是无穷多的,即顾客到达服务台后立即接受服务而无须等待.现在考虑这一随机服务系统的效率,也就是计算到时刻 t 已经服务完的顾客数与未服务完的顾客数的联合分布.

解 称某一顾客为 1 类的,如果到时刻 t 他已经服务完;称某一顾客是 2 类的,如果到时刻 t 他尚未服务完毕. $N_i(t)$ 表示到时刻 t,i 类顾客的人数,$i=1,2$.

如果顾客于时刻 $s(s\leqslant t)$ 到达,那么当他的服务时间 $Y<t-s$ 时,他是 1 类的,而服务时间的分布是 G,因此属于 1 类的概率为

$$p(s)=G(t-s)$$

于是,$N_1(t)$ 的分布是均值为

$$E[N_1(t)]=\lambda tp=\lambda\int_0^t G(t-s)\,ds=\lambda\int_0^t G(y)\,dy$$

的 Poisson 分布.

类似地,$N_2(t)$ 的分布是均值为

$$E[N_2(t)]=\lambda t(1-p)=\lambda\int_0^t[1-G(y)]\,dy$$

的 Poisson 分布,且 $N_1(t)$ 与 $N_2(t)$ 是相互独立的.

3.5 Poisson过程的推广

在前几节讨论中,可以看到Poisson过程有许多独特的性质,这与它的定义中加了许多严格限制有关,而在许多实际问题中并不都满足这些条件,因此有必要将定义中的某些条件放宽,将其推广到其他情形.

3.5.1 非齐次Poisson过程

当Poisson过程的强度λ不再是常数,而与时间t有关时,Poisson过程被推广成为非齐次(非时齐)Poisson过程. 在实际中,非齐次Poisson过程也是比较常用的. 例如在考虑设备故障率时,由于设备使用年限的变化,出故障的可能性会随之变化;放射性物质的衰变速率,会因各种外部条件的变化而随之变化;昆虫产卵的平均数量随年龄和季节的变化而变化,等等.

定义3.4 计数过程$\{N(t),t\geqslant 0\}$,$N(0)=0$,称为具有强度函数$\{\lambda(t)>0,t\geqslant 0\}$的**非齐次Poisson过程**,若满足:

(1) 独立增量过程;

(2) 对任意$t>0$和充分小的$\Delta t>0$,有

$$\begin{cases}P\{N(t+\Delta t)-N(t)=1\}=\lambda(t)\Delta t+o(\Delta t)\\P\{N(t+\Delta t)-N(t)\geqslant 2\}=o(\Delta t)\end{cases}$$

其中$\lambda(t)>0$(称为强度函数).

记$m(t)=\int_0^t\lambda(s)\mathrm{d}s$,则有

定理3.7 若$\{N(t),t\geqslant 0\}$为具有强度函数$\{\lambda(t)>0,t\geqslant 0\}$的非时齐Poisson过程,则$\forall s,t>0$,有

$$P\{N(s+t)-N(s)=n\}=\frac{[m(s+t)-m(s)]^n}{n!}\mathrm{e}^{-[m(s+t)-m(s)]}\ (n\geqslant 0)$$

非齐次Poisson过程的重要性在于不再要求平稳增量性,从而允许事件在某个时刻发生的可能性大于另一个时刻发生的可能性.

如果令$\lambda(t)\equiv\lambda$,则非齐次Poisson过程即为齐次Poisson过程. 下面的定理给出了两者之间的转换关系. 事实上,非齐次Poisson过程不过是"换了一个时钟来计时"的齐次Poisson过程.

定理3.8 (1) 设$\{N(t),t\geqslant 0\}$为具有强度函数$\{\lambda(t)>0,t\geqslant 0\}$的非时齐Poisson过程，记$M(u)=N[m^{-1}(u)]$，则$\{M(u),u\geqslant 0\}$是强度为1的时齐Poisson过程.

(2) 设$\{M(u),u\geqslant 0\}$是时齐Poisson过程，参数$\lambda=1$. 若给定强度函数$\{\lambda(s)>0,s\geqslant 0\}$，令$m(t)=\int_0^t\lambda(s)\mathrm{d}s$，$N(t)=M[m(t)]$，则$\{N(t),t\geqslant 0\}$是非时齐的具有强度函数$\{\lambda(s)>0,s\geqslant 0\}$的Poisson过程.

证明

(1) 只要证明$\{M(u),u\geqslant 0\}$满足：对任意$u>0$和充分小的$\Delta u>0$，有

$$\begin{cases}P\{M(u+\Delta u)-M(u)=1\}=\Delta u+o(\Delta u)\\P\{M(u+\Delta u)-M(u)\geqslant 2\}=o(\Delta u)\end{cases}$$

由于强度函数$\lambda(t)>0,t\geqslant 0$，而$m(t)=\int_0^t\lambda(s)\mathrm{d}s$，所以$m(t)>0$且单调递增，因此，反函数$m^{-1}(t)$存在且单调递增. 设$t=m^{-1}(u)$，$t+\Delta t=m^{-1}(u+\Delta u)$，则

$$\begin{aligned}\Delta u&=m(t+\Delta t)-m(t)=\int_t^{t+\Delta t}\lambda(s)\mathrm{d}s\\&=\lambda(t)\Delta t+o(\Delta t)\end{aligned}$$

于是

$$\begin{aligned}&\lim_{\Delta u\to 0+}\frac{P\{M(u+\Delta u)-M(u)=1\}}{\Delta u}\\=&\lim_{\Delta t\to 0+}\frac{P\{N[m^{-1}(u+\Delta u)]-N[m^{-1}(u)]=1\}}{\lambda(t)\Delta t+o(\Delta t)}\\=&\lim_{\Delta t\to 0+}\frac{P\{N(t+\Delta t)-N(t)=1\}}{\lambda(t)\Delta t+o(\Delta t)}\\=&\lim_{\Delta t\to 0+}\frac{\lambda(t)\Delta t+o(\Delta t)}{\lambda(t)\Delta t+o(\Delta t)}=1\end{aligned}$$

即

$$P\{M(u+\Delta u)-M(u)=1\}=\Delta u+o(\Delta u)$$

同理可得

$$P\{M(u+\Delta u)-M(u)\geqslant 2\}=o(\Delta u)$$

所以$\{M(u),u\geqslant 0\}$是强度参数为1的时齐Poisson过程.

(2) 证略.

【例3.13】 设某设备的使用期限为10年，在前5年内平均2.5年需要维修一次，后5年平均2年需要维修一次. 试求它在使用期限内只维修过一次的概率.

解 $N(t)$表示$(0,t]$内该设备维修的次数，可以用非齐次Poisson过程来描述该过程，其强度函数为

$$\lambda(t)=\begin{cases}1/2.5, & 0\leqslant t\leqslant 5\\ 1/2, & 5<t\leqslant 10\end{cases}$$

$$m(10)=\int_0^{10}\lambda(t)\mathrm{d}t=\int_0^5\frac{1}{2.5}\mathrm{d}t+\int_5^{10}\frac{1}{2}\mathrm{d}t=4.5$$

故在使用期限内只维修过一次的概率为

$$P\{N(10)-N(0)=1\}=\mathrm{e}^{-4.5}\frac{(4.5)^1}{1!}=\frac{9}{2}\mathrm{e}^{-\frac{9}{2}}$$

【例3.14】 从历史观测可以发现，平日上午5：00～11：00到某个加油站加油的汽车数量可以用非齐次Poisson过程$\{N(t),t\geqslant 0\}$来模拟，并且具有以下强度函数：

$$\lambda(t)=10+35.4(t-5)\mathrm{e}^{-(t-5)^2/8},5\leqslant t\leqslant 11$$

问：

(1) 在某个工作日上午5：00～11：00来加油的汽车数量的均值是多少？

(2) 至少有90辆车在某个工作日的6：00～8：00来加油的概率是多少？

解 (1) 上午5：00～11：00来加油的汽车数量的均值为

$$E[N(5,11)]=\int_5^{11}\lambda(t)\mathrm{d}t=\int_0^6(10+35.4t\mathrm{e}^{-t^2/8})\mathrm{d}t$$
$$=[10t-141.6\mathrm{e}^{-t^2/8}]_0^6=200$$

(2) 在6：00～8：00到达车辆的均值为

$$\int_6^8\lambda(t)\mathrm{d}t=\int_1^3(10+35.4t\mathrm{e}^{-t^2/8})\mathrm{d}t$$
$$=[10t-141.6\mathrm{e}^{-t^2/8}]_1^3=99$$

因此，在6：00～8：00到达的车辆数$N(6,8)=N(8)-N(6)$具有参数为99的Poisson分布，即

$$P\{N(6,8)\geqslant 90\}=\sum_{n=90}^{+\infty}\frac{99^n}{n!}\mathrm{e}^{-0.99}$$

通过Poisson分布的正态近似，得到

$$\sum_{n=90}^{+\infty}\frac{99^n}{n!}\mathrm{e}^{-0.99}\approx 1-\Phi\left(\frac{90-99}{\sqrt{99}}\right)\approx 1-0.1827=0.8173$$

因此，$P\{N(6,8)\geqslant 90\}=0.8173$.

【例3.15】 某小商店上午8时开始营业，从8时到11时平均顾客到达率呈线性增加；从8时开始顾客到达率为5人/h，11时到达率达到高峰，为20人/h；从11时至下午1时到达率不变；从下午1时至5时顾客到达率呈线性下降，到下午5时顾客到达率为12人/h. 设在不相重叠的时间间隔内到达的顾客数是相互独立的. 试求在上午8时半至9时半无顾客到达的概率和该段时间内到达顾客数的数学期望.

解 因为顾客到达率与时间有关，所以顾客到达过程是一个非齐次 Poisson 过程.

设上午 8 时为 $t=0$，则上午 11 时为 $t=3$，下午 1 时为 $t=5$，下午 5 时为 $t=9$. 次日 8 时从 $t=0$ 重新开始，故可得一周期为 9 的曲线. 按曲线写出顾客到达率 $\lambda(t)$ 的关系式如下：

$$\lambda(t)=\begin{cases}5+5t, & 0\leqslant t\leqslant 3\\ 20, & 3<t\leqslant 5;\\ 20-2(t-5), & 5<t\leqslant 9\end{cases}\quad 且\ \lambda(t)=\lambda(t+9)$$

因为

$$m\left(\frac{3}{2}\right)-m\left(\frac{1}{2}\right)=\int_{1/2}^{3/2}\lambda(t)\mathrm{d}t=\int_{1/2}^{3/2}(5+5t)\mathrm{d}t=10$$

所以在上午 8 时半至 9 时半无顾客到达的概率为 $\mathrm{e}^{-[m(3/2)-m(1/2)]}=\mathrm{e}^{-10}$，在上午 8 时半至 9 时半的时间内到达顾客数的数学期望为 $m(3/2)-m(1/2)=10$(人).

3.5.2 复合 Poisson 过程

人们在考虑设备故障所需的维修费，自然灾害所造成的损失，股票市场的价格变动时，都会碰到这样一类模型：事件的发生服从 Poisson 过程，而每一次事件都还附带一个随机变量(如费用、损失等). 此时人们感兴趣的不仅仅是事件发生的次数，还需要了解总费用或总损失等. 这时就需要用复合 Poisson 过程来描述这一过程.

定义 3.5 设$\{Y_k,k\geqslant 1\}$是独立同分布的随机变量序列，$\{N(t),t\geqslant 0\}$为 Poisson 过程，且$\{N(t),t\geqslant 0\}$与$\{Y_i,i\geqslant 1\}$独立，记

$$X(t)=\sum_{i=1}^{N(t)}Y_i$$

则称$\{X(t),t\geqslant 0\}$为复合 Poisson 过程.

若 $N(t)$ 表示$[0,t]$内到达的粒子数，Y_i 表示第 i 个粒子的能量，则 $X(t)$ 表示$[0,t]$内到达的粒子的总能量；若$\{N(t),t\geqslant 0\}$表示顾客流，Y_i 表示第 i 个顾客的行李重量，则 $X(t)$ 表示$[0,t]$内到达的顾客的行李总重量；若在某保险公司买了人寿保险的人在时刻 $W_1,W_2,\cdots,W_n,\cdots$ 死亡，在时刻 W_n 死亡的人的保险金额是 Y_n，在$[0,t]$内死亡的人数为 $N(t)$，则 $X(t)=\sum_{i=1}^{N(t)}Y_i$ 表示该公司在$[0,t]$内需要支付的赔偿金总额.

容易看出，复合 Poisson 过程不一定是计数过程，但当 $Y_i\equiv c,i=1,2,\cdots,c$ 为常数时，可转化为 Poisson 过程. 对于复合 Poisson 过程，通常关心的是它的一些数

字特征.

定理 3.9　设$\{X(t)=\sum_{i=1}^{N(t)}Y_i,t\geqslant 0\}$是复合Poisson过程,$\{N(t),t\geqslant 0\}$是强度为$\lambda$的Poisson过程,则

(1)$X(t)$有独立增量;

(2)若$E(Y_i^2)<+\infty$,则

$$E[X(t)]=\lambda tE(Y_1);\quad D[X(t)]=\lambda tE(Y_1^2)$$

证明　(1)令$0\leqslant t_0<t_1<t_2<\cdots<t_n$,则

$$X(t_k)-X(t_{k-1})=\sum_{i=N(t_{k-1})+1}^{N(t_k)}Y_i,\ k=1,2,\cdots,n$$

由Poisson过程的独立增量性和Y_i之间的独立性可知$X(t)$的独立增量性.

(2)

$$\begin{aligned}\varphi_{X(t)}(u)&=E[\mathrm{e}^{uX(t)}]=\sum_{n=0}^{+\infty}E[\mathrm{e}^{uX(t)}\mid N(t)=n]P\{N(t)=n\}\\&=\sum_{n=0}^{+\infty}E[\mathrm{e}^{u(Y_1+Y_2+\cdots+Y_n)}\mid N(t)=n]\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}\\&=\sum_{n=0}^{+\infty}E[\mathrm{e}^{u(Y_1+Y_2+\cdots+Y_n)}]\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}\\&=\sum_{n=0}^{+\infty}[E(\mathrm{e}^{uY_1})]^n\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}\end{aligned}$$

令$Y_i\sim Y$的矩母函数为$\varphi_Y(u)=E(\mathrm{e}^{uY})$,则有

$$\varphi_{X(t)}(u)=E[\mathrm{e}^{uX(t)}]=\sum_{n=0}^{+\infty}\frac{[\lambda t\varphi_Y(u)]^n}{n!}\mathrm{e}^{-\lambda t}=\exp\{\lambda t[\varphi_Y(u)-1]\}$$

对上式在$u=0$处求导数,有

$$E[X(t)]=\varphi'_{X(t)}(0)=\lambda tE(Y_1)$$

以及

$$D[X(t)]=\lambda tE(Y_1^2)$$

【例 3.16】　上海证券交易所,宝钢股份的交易流是强度为λ(笔/min)的Poisson过程.设第j笔交易量是Z_j手,如果$\{Z_j\}$是来自总体Z的随机变量,且$\mu=E(Z)$,$\sigma^2=\mathrm{Var}(Z)$.计算宝钢股份60 min内的交易量的数学期望和标准差.

解　用$\{N(t),t>0\}$表示上述的强度为λ的Poisson过程.60 min的交易量为$X(60)=\sum_{j=1}^{N(60)}Z_j$,根据定理3.9知,60 min的平均交易量为

$$E[X(60)]=60\lambda\mu(\text{手})$$

交易量的标准差为

$$\sqrt{\mathrm{Var}[X(60)]} = \sqrt{60\lambda(\mu^2+\sigma^2)} = \sqrt{60\lambda E(Z^2)}(\text{手})$$

【例 3.17】 保险公司接到的索赔次数服从强度为每月 2 次的 Poisson 过程 $\{N(t),t\geqslant 0\}$，每次要求赔偿的金额 Y_i 相互独立同分布，服从均值为 10000 的正态分布，则一年中保险公司平均的赔付额是多少?

解 在$[0,t]$内保险公司需要赔付的总金额$\{X(t)=\sum_{i=1}^{N(t)}Y_i,t\geqslant 0\}$是一个复合 Poisson 过程，所以一年中保险公司平均的赔付额为

$$E[X(12)] = 2\times 12\times 10000 = 240000$$

【例 3.18】 设移民到某地定居的户数是 Poisson 过程，平均每周有 2 户定居. 设每户的人口数是一随机变量，一户有 4 人的概率是 1/6，有 3 人的概率是 1/3，有 2 人的概率是 1/3，有 1 人的概率是 1/6. 求在五周内到该地定居的移民人数的数学期望和方差.

解 以 Y_i 记第 i 户人口数，$N(t)$ 表示$(0,t]$内到该地定居的户数，则移民总人数 $X(t)=\sum_{i=1}^{N(t)}Y_i$ 是一复合 Poisson 过程.

由题意知 $\lambda=2,t=5$，而

$$E(Y_1) = 4\times\frac{1}{6}+3\times\frac{1}{3}+2\times\frac{1}{3}+1\times\frac{1}{6}=\frac{5}{2}$$

$$E(Y_1^2) = 4^2\times\frac{1}{6}+3^2\times\frac{1}{3}+2^2\times\frac{1}{3}+1^2\times\frac{1}{6}=\frac{43}{6}$$

所以 $E[X(5)]=2\times 5\times\frac{5}{2}=25$，$D[X(5)]=2\times 5\times\frac{43}{6}=\frac{215}{3}$.

3.5.3 条件 Poisson 过程

Poisson 过程描述的是一个有着“风险”参数 λ 的个体发生某一事件的概率. 如果考虑一个总体，其中的个体存在差异，比如发生事故的倾向性因人而异，这时可以把定义 3.2 中的概率分布解释为给定 λ 时，$N(t)$ 的条件分布.

定义 3.6 设 Λ 是一正的随机变量，分布函数为 $G(x)$，$x\geqslant 0$，设$\{N(t),t\geqslant 0\}$是一计数过程，若在给定条件 $\Lambda=\lambda$ 下，$\{N(t),t\geqslant 0\}$是一 Poisson 过程，即 $\forall s,t\geqslant 0,n\in N_0,\lambda\geqslant 0$，有

$$P\{N(s+t)-N(s)=n\mid\Lambda=\lambda\}=\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}$$

则称$\{N(t),t\geqslant 0\}$是条件 Poisson 过程，记为$\{N_\Lambda(t),t\geqslant 0\}$.

定理 3.10 设$\{N_\Lambda(t),t\geqslant 0\}$是条件 Poisson 过程，且 $E(\Lambda^2)<+\infty$，则

(1)$P\{N(s+t)-N(s)=n\}=\int_0^{+\infty}\mathrm{e}^{-\lambda t}\dfrac{(\lambda t)^n}{n!}\mathrm{d}G(\lambda)$;

(2)$E[N_\Lambda(t)]=tE(\Lambda)$;

(3)$E[N_\Lambda(t)^2]=t^2E(\Lambda)^2+tE(\Lambda)$.

证明

(1) 由全概率公式知，若 $\Lambda\sim G(\lambda)$，$\lambda>0$，且

$$P\{N(s+t)-N(s)=n\mid\Lambda=\lambda\}=\frac{(\lambda t)^n\mathrm{e}^{-\lambda t}}{n!}$$

则

$$\begin{aligned}P\{N(s+t)-N(s)=n\}&=\int_0^{+\infty}P\{N(s+t)-N(s)=n\mid\Lambda=\lambda\}\mathrm{d}G(\lambda)\\&=\int_0^{+\infty}\mathrm{e}^{-\lambda t}\frac{(\lambda t)^n}{n!}\mathrm{d}G(\lambda)\end{aligned}$$

(2) 由期望的全概率公式得

$$E[N_\Lambda(t)]=\int_0^{+\infty}E[N(t)\mid\Lambda=\lambda]\mathrm{d}G(\lambda)=\int_0^{+\infty}\lambda t\,\mathrm{d}G(\lambda)=tE(\Lambda)$$

(3)
$$\begin{aligned}E[N_\Lambda(t)^2]&=\int_0^{+\infty}E[N(t)^2\mid\Lambda=\lambda]\mathrm{d}G(\lambda)\\&=\int_0^{+\infty}[(\lambda t)^2+\lambda t]\mathrm{d}G(\lambda)=t^2E(\Lambda^2)+tE(\Lambda)\end{aligned}$$

下面的定理列出了条件 Poisson 过程的两个重要性质.

定理 3.11 (1) 条件 Poisson 过程$\{N_\Lambda(t),t\geqslant 0\}$具有平稳增量；

(2) 如果 Λ 不是常量，那么条件 Poisson 过程$\{N_\Lambda(t),t\geqslant 0\}$的增量是非独立的.

证明

(1) 假设 $0=t_0<t_1<\cdots<t_n$；$n=1,2,\cdots$，那么，对于任何非负整数 $i_1,i_2,\cdots,i_n$，有

$$\begin{aligned}&P\{N_\Lambda(t_{k-1}+\tau,t_k+\tau)=i_k;k=1,2,\cdots,n\}\\&=\int_0^{+\infty}P\{N_\lambda(t_{k-1}+\tau,t_k+\tau)=i_k;k=1,2,\cdots,n\}\mathrm{d}G(\lambda)\\&=\int_0^{+\infty}P\{N_\lambda(t_{k-1},t_k)=i_k;k=1,2,\cdots,n\}\mathrm{d}G(\lambda)\\&=P\{N_\Lambda(t_{k-1},t_k)=i_k;k=1,2,\cdots,n\}.\end{aligned}$$

(2) 假设 $0\leqslant t_1<t_2<t_3$，那么

$$\begin{aligned}
&P\{N_\Lambda(t_1,t_2)=i_1,N_\Lambda(t_2,t_3)=i_2\}\\
&=\int_0^{+\infty}P\{N_\lambda(t_1,t_2)=i_1,N_\lambda(t_2,t_3)=i_2\}\mathrm{d}G(\lambda)\\
&=\int_0^{+\infty}P\{N_\lambda(t_1,t_2)=i_1\}P\{N_\lambda(t_2,t_3)=i_2\}\mathrm{d}G(\lambda)\\
&\neq\int_0^{+\infty}P\{N_\lambda(t_1,t_2)=i_1\}\mathrm{d}G(\lambda)\int_0^{+\infty}P\{N_\lambda(t_2,t_3)=i_2\}\mathrm{d}G(\lambda)\\
&=P\{N_\Lambda(t_1,t_2)=i_1\}P\{N_\Lambda(t_2,t_3)=i_2\}
\end{aligned}$$

习　题　3

1. 在某地的灾难性车祸可以用强度为 $\lambda=3$(次 / 年) 的齐次 Poisson 过程描述,则在最近一年的下半年至少有 2 次灾难性车祸发生的概率是多少?
2. $\{N(t),t\geqslant 0\}$是强度为 λ 的齐次 Poisson 过程,证明:其协方差函数 $C_X(s,t)=\lambda\min(s,t)$.
3. 每天 12∶00 ～ 14∶00 通过某十字路口的车辆,服从强度为 $\lambda=40$(辆 /h) 的齐次 Poisson 过程,其中有 0.8% 的车辆会忽视指示牌"停". 问在 12∶00 ～ 13∶00 间至少有一辆汽车忽视指示牌的概率是多大?
4. 设$\{N(t),t\geqslant 0\}$是一个强度为λ的齐次Poisson过程,对于一个任意固定的正数 h,令 $X(t)=N(t+h)-N(t)$,证明:随机过程$\{X(t),t\geqslant 0\}$是弱平稳的.
5. 大量样本的统计评价模型证明,每天 12∶00 ～ 16∶00 之间到达某汽车加油站的汽车数量$\{N(t),t\geqslant 0\}$为非齐次 Poisson 过程,强度为
$$\lambda(t)=8-4t+3t^2,0\leqslant t\leqslant 4$$
求:
(1) 在 12∶00 ～ 16∶00 之间平均有多少辆汽车到达加油站?
(2) 在 12∶00 ～ 16∶00 之间至少有 40 辆汽车达到的概率.
6. 设$\{N(t),t\geqslant 0\}$ 为 Poisson 过程,参数为 λ,求或证明:
(1) $E\{N(t)N(s+t)\}$;
(2)$E[N(t+s)\mid N(s)]$ 的分布律;
(3) 任给 $0\leqslant s\leqslant t$, 有 $P\{N(s)\leqslant N(t)\}=1$;
(4) 任给 $0\leqslant s\leqslant t$, $\varepsilon>0$,有$\lim\limits_{t\to s}P\{N(t)-N(s)>\varepsilon\}=0$.
7. 设$\{N(t),t\geqslant 0\}$是具有参数为λ的Poisson过程,假定T是相邻事件的时间间隔,证明:
$$P\{T>t_1+t_2\mid T>t_1\}=P\{T>t_2\}$$
8. 设 $N_1(t)$ 和 $N_2(t)$ 是强度分别为 λ_1 和 λ_2 且相互独立的 Possion 过程. 证明:在 $N_1(t)$ 的任意两个相邻事件之间的时间间隔内,$N_2(t)$ 恰好有 k 个事件发生的概

率为

$$p = \frac{\lambda_1}{\lambda_1 + \lambda_2}\left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^k$$

9. 设 $X, Y_1, Y_2, \cdots$ 是一列相互独立的随机变量，其中 X 服从参数为 λ 的 Poisson 分布，$Y_1, Y_2, \cdots$ 服从 $[0,1]$ 上的均匀分布，定义

$$N(t) = \sum_{k=1}^{X} I_{[0,t]}(Y_k), t \in [0,1]$$

其中 $I_{[0,t]}(\cdot)$ 表示 $[0,t]$ 上的示性函数，证明：$\{N(t), t \in [0,1]\}$ 是参数为 λ 的 Poisson 过程.

10. 在一个二手车经销商店，到达的特定类型的汽车数量 $\{N(t), t \geqslant 0\}$ 服从强度为 λ 的齐次 Poisson 过程. 设 $T_i, i = 1,2,\cdots$ 是该类型车第 i 辆的到达时间，在 T_i 时刻到达的汽车马上就会被经销商以价格 C_i 转售，$C_1, C_2, \cdots$ 相互独立且同分布. 但是如果买家在 $T_i + \tau$ 时刻买在 T_i 时刻到达的车，他只需要付 $\mathrm{e}^{-\alpha\tau} C_i$，其中 $\alpha > 0$. 在 t 时刻，经销商同时出售所有类型的汽车. 求汽车经销商出售汽车总价格的平均值.

11. $\{N(t), t \geqslant 0\}$ 是强度函数为 $\lambda(t)$ 的非齐次 Poisson 过程，均值函数 $\Lambda(t) = \int_0^t \lambda(x)\mathrm{d}x$，记第 i 次事件发生的时刻为 T_i，证明：在 $N(t) = n$ 的条件下，随机向量 $(T_1, T_2, \cdots, T_n)$ 与 n 个独立同分布，分布函数为

$$F(x) = \begin{cases} \dfrac{\Lambda(x)}{\Lambda(t)}, & 0 \leqslant x \leqslant t \\ 1, & x > t \end{cases}$$

的随机变量的顺序统计量的分布相同.

12. 设 $\{N(t), t \geqslant 0\}$ 是一个强度为 λ 的齐次 Poisson 过程，令 $T_i, i = 1,2,\cdots$ 是第 i 次事件发生的时刻，确定过程 $\{X(t), t \geqslant 0\}$，

$$X(t) = \sum_{i=1}^{N(t)} h(t - T_i), \text{其中 } h(t) = \begin{cases} \sin t, & 0 \leqslant t \leqslant \pi \\ 0, & \text{其他} \end{cases}$$

的协方差函数 $C_X(\tau)$.

第4章　马尔可夫链

4.1　马尔可夫过程的概念

马尔可夫(Markov)过程主要应用于排队论、存储模型和更新模型方面，也广泛应用于网络流量分析和计算机系统建模等方面．特别是近期Markov过程的计算机仿真算法理论的快速发展，使得Markov过程的应用取得了许多革新成果．

定义4.1　如果对于T中任意有限个点$t_1 < t_2 < \cdots < t_n$，以及E中任意点$x_i(i = 1,2,\cdots,n)$，x和E中的任意波雷尔(Borel)集A，都有

$$P\{X(t) \in A \mid X(t_1) = x_1, X(t_2) = x_2, \cdots, X(t_n) = x_n, X(t) = x\}$$
$$= P\{X(t) \in A \mid X(t) = x\}$$

则称随机过程$\{X(t), t \in T\}$为Markov过程，称其条件具有Markov性．

上述结论也可写成

$$P\{X(t_n) \leqslant x_n \mid X(t_1) = x_1, X(t_2) = x_2, \cdots, X(t_{n-1}) = x_{n-1}\}$$
$$= P\{X(t_n) \leqslant x_n \mid X(t_{n-1}) = x_{n-1}\},\ x_n \in \mathbf{R}$$

或 $F_{t_n \mid t_1 t_2 \cdots t_{n-1}}(x_n, t_n \mid x_1, x_2, \cdots, x_{n-1}; t_1, t_2, \cdots, t_{n-1}) = F_{t_n \mid t_{n-1}}(x_n, t_n \mid x_{n-1}, t_{n-1})$

或 $f_{t_n \mid t_1 t_2 \cdots t_{n-1}}(x_n, t_n \mid x_1, x_2, \cdots, x_{n-1}; t_1, t_2, \cdots, t_{n-1}) = f_{t_n \mid t_{n-1}}(x_n, t_n \mid x_{n-1}, t_{n-1})$

参数集$T = (0, +\infty)$或$T = \{0,1,2,\cdots\} = \mathbf{N}^+$，状态空间$E = \mathbf{R}$或$\mathbf{R}$中的某波雷尔集．

由定义4.1可知，Markov过程具有Markov性或无后效性，即在过程或系统在时刻t_0所处的状态为已知的条件下，过程在时刻$t(t > t_0)$所处的状态的条件分布与过程在时刻t_0之前所处的状态无关．也就是说，在已知过程的“现在”的条件下，过程的“将来”与过程的“过去”无关，只与“现在”有关．

【例4.1】　设$\{X(t), t \in T\}$是一独立随机过程，即对于n个$t_1, t_2, \cdots, t_n \in T$，$X(t_1), X(t_2), \cdots, X(t_n)$总体独立，证明：$\{X(t), t \in T\}$是一马尔可夫过程．

证明　由已知条件知，随机事件

$X(t_1) = x_1, X(t_2) = x_2, \cdots, X(t_n) = x_n, X(t) \leqslant x$也相互独立，于是

$$P\{X(t) \leqslant x \mid X(t_1) = x_1, X(t_2) = x_2, \cdots, X(t_n) = x_n\} = P\{X(t) \leqslant x\}$$
$$P\{X(t) \leqslant x \mid X(t_n) = x_n\} = P\{X(t) \leqslant x\}$$

从而 $P\{X(t)\leqslant x \mid X(t_1)=x_1,X(t_2)=x_2,\cdots,X(t_n)=x_n\}$

$$=P\{X(t)\leqslant x \mid X(t_n)=x_n\}$$

由定义 4.1 知，$\{X(t),t\in T\}$ 是马尔可夫过程.

同理也可以证明独立随机序列的和、常用的泊松过程以及维纳(Wiener)过程等也都是马尔可夫过程.

Markov 过程按其状态和时间参数是连续的或离散的，可以分为如下四类：

(1) 时间参数离散、状态空间离散的过程，称为马尔可夫链，简称马氏链.

(2) 时间参数连续、状态空间离散的过程，称为纯不连续 Markov 过程或可数状态的 Markov 过程.

(3) 时间参数离散、状态空间连续的过程，称为 Markov 序列.

(4) 时间参数连续、状态空间连续的过程，称为连续马尔可夫过程或扩散过程.

由于课时数的限制，本章将主要介绍 Markov 链的一些基础理论知识及其在相关领域内的某些应用.

4.2 马尔可夫链的概念

4.2.1 马尔可夫链的定义

设马尔可夫过程 $\{X_n;n\in T\}$ 的参数集 T 是离散的时间集合，即 $T=\{0,1,2,\cdots\}$，其相应 X_n 可能取值的全体组成的状态空间是离散的状态集 $I=\{i_0,i_1,i_2,\cdots\}$.

定义 4.2 设随机序列 $\{X_n;n\geqslant 0\}$ 的状态空间为 I，如果对 $\forall n\in N_0$，及 $i_0,i_1,\cdots,i_n,i_{n+1}\in I$，$P\{X_0=i_0,X_1=i_1,\cdots,X_n=i_n\}>0$，有

$$P\{X_{n+1}=i_{n+1}\mid X_0=i_0,X_1=i_1,\cdots,X_n=i_n\}=P\{X_{n+1}=i_{n+1}\mid X_n=i_n\} \tag{4.1}$$

则称 $\{X_n;n\geqslant 0\}$ 为 **Markov 链**.

可以看出，式(4.1)有效地刻画了 Markov 链的特性，称此特性为 Markov 性或无后效性，简称为马氏性. 简言之，即已知“现在”，“将来”与“过去”无关. Markov 链也称为马氏链.

由定义 4.2 知

$$P\{X_0=i_0,X_1=i_1,\cdots,X_n=i_n\}$$

$$
\begin{aligned}
&= P\{X_n = i_n \mid X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}\} \times \\
&\quad P\{X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}\} \\
&= P\{X_n = i_n \mid X_{n-1} = i_{n-1}\} \times P\{X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}\} \\
&= P\{X_n = i_n \mid X_{n-1} = i_{n-1}\} \times P\{X_{n-1} = i_{n-1} \mid X_{n-2} = i_{n-2}\} \times \cdots \times \\
&\quad P\{X_1 = i_1 \mid X_0 = i_0\} \times P\{X_0 = i_0\}.
\end{aligned}
$$

可见，如果马尔可夫链的初始分布 $P\{X_0 = i_0\}$ 确定，其统计性质完全由条件概率 $P\{X_{n+1} = i_{n+1} \mid X_n = i_n\}$ 决定. 如何确定这个条件概率，是马尔可夫链理论与应用研究中的重要问题之一.

4.2.2　转移概率

定义 4.3　设 $\{X_n; n \geqslant 0\}$ 为马氏链，状态空间为 I，对于 $\forall i,j \in I$，称

$$P\{X_{n+1} = j \mid X_n = i\} = p_{ij}(n)$$

为马氏链 $\{X_n; n \geqslant 0\}$ 在 n 时刻的**一步转移概率**，简称为**转移概率**.

显然，一步转移概率满足：

$$p_{ij}(n) \geqslant 0, \quad i,j \in I$$

$$\sum_{j \in I} p_{ij}(n) = 1, i \in I$$

一般地，转移概率 $p_{ij}(n)$ 不仅与状态 i,j 有关，而且与时刻 n 有关. 当 $p_{ij}(n)$ 不依赖于时刻 n 时，表示马尔可夫链具有平稳转移概率.

若对于 $\forall i,j \in I$，有

$$P\{X_{n+1} = j \mid X_n = i\} = p_{ij}(n) \equiv p_{ij}$$

即上面式子的右边与时刻 n 无关，则称此马氏链为**齐次(或时齐的) 马氏链**.

如果没有特别说明，下面只讨论齐次马尔可夫链，通常将“齐次”两字省略.

设 $p_0(i) = P\{X_0 = i\}, i \in I$，如果对一切 $i \in I$ 都有 $p_0(i) \geqslant 0$，$\sum_{i \in I} p_0(i) = 1$，则称 $p_0(i)$ 为马氏链的**初始分布**. 记

$$\pi_i(n) = P\{X_n = i\}, \ \boldsymbol{\pi}(n) = (\pi_1(n), \pi_2(n), \cdots, \pi_i(n), \cdots)$$

$\boldsymbol{\pi}(n)$ 表示 n 时刻 X_n 的概率分布向量. 称 $\{\pi_i(n), i \in I\}$ 为 **Markov 链的绝对分布**；$\{\pi_i(0), i \in I\}$ 为 Markov 链的初始分布.

若绝对分布 $\pi_i(n)$ 与 n 无关，即

$$\pi_i(n) = P\{X_n = i\} = \pi_i, i \in I, n \geqslant 0$$

则称 $\{\pi_i, i \in I\}$ 为马氏链 $\{X_n; n \geqslant 0\}$ 的**定态分布**.

对于齐次马氏链，记 $\boldsymbol{P} = (p_{ij})$，且状态空间 $I = \{1,2,\cdots\}$，则称矩阵 $\boldsymbol{P}$ 为马氏链的**一步转移概率矩阵**，简称为**转移矩阵**.

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} & \cdots \\ p_{21} & p_{22} & \cdots & p_{2n} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

它具有以下性质：

(1) $p_{ij} \geqslant 0, i, j \in I$;

(2) $\sum_{j \in I} p_{ij} = 1, i \in I$.

通常满足上述性质(1) 和(2) 的矩阵，称为**随机矩阵**.

一个 Markov 链的特性完全由它的一步转移概率矩阵 $\boldsymbol{P}$ 及初始分布向量 $\{\pi_i(0), i \in I\}$ 决定.

定理 4.1

$$\boldsymbol{\pi}(n+1) = \boldsymbol{\pi}(n)\boldsymbol{P} \tag{4.2}$$

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\boldsymbol{P}^n \tag{4.3}$$

证明 先对事件进行分解

$$\{X_{n+1} = j\} = \bigcup_{i \in I} \{X_n = i, X_{n+1} = j\}$$

因为当 $i \neq k$ 时，$\{X_n = i\} \cap \{X_n = k\} = \varnothing$，故

$$\begin{aligned} P\{X_{n+1} = j\} &= \sum_{i \in I} P\{X_n = i, X_{n+1} = j\} \\ &= \sum_{i \in I} P\{X_n = i\} P\{X_{n+1} = j \mid X_n = i\} = \sum_{i \in I} \pi_i(n) p_{ij} \end{aligned}$$

写成向量形式即得

$$\boldsymbol{\pi}(n+1) = \boldsymbol{\pi}(n)\boldsymbol{P}$$

重复利用式(4.2) 即得式(4.3).

如果状态空间是有限的，设状态数为 n，则一步转移概率矩阵是 n 阶方阵；若状态是无限可列的情形，则一步转移概率矩阵只是形式上的矩阵.

【例 4.2】 设$\{Y_k, k \geqslant 1\}$是一个独立同分布的、取非负整数的随机变量序列，$P\{Y_k = i\} = a_i (i \geqslant 0)$. 令 $X_n = (\sum_{k=1}^{n} Y_k)^2, n \geqslant 1$，证明：$\{X_n, n \geqslant 0\}$ 是马尔可夫链，并求其一步转移概率矩阵.

解 X_n 的值域为$\{0, 1, 2^2, 3^2, \cdots\} = E$，$X_{n+1} = X_n + 2\sqrt{X_n} Y_{n+1} + Y_{n+1}^2, n \geqslant 1$，有

$$\begin{aligned} & P\{X_{n+1} = i_{n+1} \mid X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_1 = i_1\} \\ =\ & P\{i_n + 2\sqrt{i_n} Y_{n+1} + Y_{n+1}^2 = i_{n+1}\} \\ =\ & P\{X_{n+1} = i_{n+1} \mid X_n = i_n\} \end{aligned}$$

所以，$\{X_n, n \geqslant 0\}$ 是马尔可夫链. 其转移概率为

$$p_{ij} = P\{i + 2\sqrt{i} Y_1 + Y_1^2 = j\} = a_{\sqrt{j} - \sqrt{i}} \quad (i, j \in E, j \geqslant i)$$

$$p_{ij} = 0 \quad (j < i)$$

故，转移概率矩阵为

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & & \\ & a_0 & a_1 & a_2 & \cdots & \\ & & a_0 & a_1 & a_2 & \cdots \\ & & & \ddots & \ddots & \ddots \end{pmatrix}$$

4.2.3 m 步转移概率与 Chapman-Kolmogorov 方程

定义 4.4 称 $p_{ij}^{(m)}(n) = P\{X_{n+m} = j \mid X_n = i\}$ 为马氏链 $\{X_n, n \geqslant 0\}$ 的 m **步转移概率**. 在齐次马氏链的情况下，$p_{ij}^{(m)}(n)$ 与 n 无关，记为 $p_{ij}^{(m)}$，称

$$P^{(m)} = (p_{ij}^{(m)})$$

为齐次马氏链的 m **步转移(概率) 矩阵**.

m 步转移概率矩阵显然满足：

$$p_{ij}^{(m)}(n) \geqslant 0, \quad i,j \in I$$

$$\sum_{j\in I} p_{ij}^{(m)}(n) = 1, \quad i \in I$$

$m = 1$ 时，即为一步转移概率矩阵.

规定：

$$p_{ij}^{(0)}(n) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

定理 4.2 (C-K 方程) 对于 m 步转移概率有如下的方程：

$$p_{ij}^{(m+r)}(n) = \sum_{k\in I} p_{ik}^{(m)}(n) p_{kj}^{(r)}(n+m), \quad i,j \in I$$

对于齐次马氏链，此方程为

$$p_{ij}^{(m+r)} = \sum_{k\in I} p_{ik}^{(m)} p_{kj}^{(r)}, \quad i,j \in I \quad \text{(C-K 方程)}$$

称上述方程为 Chapman-Kolmogorov 方程，即切普曼 - 柯尔莫哥洛夫(C-K) 方程.

证明 由全概率公式，有

$$\begin{aligned} p_{ij}^{(m+r)}(n) &= P\{X_{n+m+r} = j \mid X_n = i\} \\ &= \sum_{k\in I} P\{X_{n+m+r} = j, X_{n+m} = k \mid X_n = i\} \\ &= \sum_{k\in I} P\{X_{n+m+r} = j \mid X_{n+m} = k, X_n = i\} \cdot P\{X_{n+m} = k \mid X_n = i\} \\ &= \sum_{k\in I} P\{X_{n+m+r} = j \mid X_{n+m} = k\} P\{X_{n+m} = k \mid X_n = i\} \end{aligned}$$

$$= \sum_{k \in I} p_{ik}^{(m)}(n) p_{kj}^{(r)}(n+m)$$

C-K 方程表明，从状态 i 出发经过 $m+r$ 步到达状态 j 的过程可分为两个阶段：先从状态 i 出发经过 m 步到达状态 k，再由状态 k 出发经过 r 步到达状态 j. 由马氏性可知，后一阶段的状态转移与前一阶段的状态转移相互独立，故两阶段的转移概率是相乘的关系. 而经过 m 步到达状态 k 不受任何限制，因此要对全部的 k 求和.

由 C-K 方程易得如下结果：

推论 4.1 $p_{ij}^{(k+1)}(n) = \sum_{s_1 \in I} \sum_{s_2 \in I} \cdots \sum_{s_k \in I} p_{is_1}(n) p_{s_1 s_2}(n+1) \cdots p_{s_k j}(n+k)$

推论 4.2 $p_{ij}^{(k+1)}(n) = \sum_{r \in I} p_{ir}(n) p_{rj}(n+1) = \sum_{r \in I} p_{ir}^{(k)}(n) p_{rj}(n+k)$

对于齐次马氏链的情形，可以把 C-K 方程写成矩阵的形式，即有

$$\boldsymbol{P}^{(m+r)} = \boldsymbol{P}^{(m)} \boldsymbol{P}^{(r)}$$

由此推出

$$\boldsymbol{P}^{(m)} = \boldsymbol{P}^{(m-1)} \boldsymbol{P}^{(1)} = \cdots = (\boldsymbol{P})^m = \boldsymbol{P}^m$$

其中：$\boldsymbol{P}^{(1)} = \boldsymbol{P}$.

4.2.4 Markov 链的基本性质

由上述定理和结论不难得到如下马尔可夫链的基本性质：

性质 1 设 $\{X_n, n \geqslant 0\}$ 为马氏链，其状态空间为 I，则

$$\begin{aligned} & P\{X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n\} \\ = & P\{X_0 = i_0\} P\{X_1 = i_1 \mid X_0 = i_0\} \\ & \times P\{X_2 = i_2 \mid X_1 = i_1\} \cdots P\{X_n = i_n \mid X_{n-1} = i_{n-1}\} \end{aligned}$$

性质1表明，$X_0, X_1, X_2, \cdots, X_n$ 的有限维联合分布可由初始分布及转移概率所决定，即有

$$P\{X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n\} = p_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}$$

性质 2 设 $\{X_n, n \geqslant 0\}$ 为马氏链，其状态空间为 I，则

$$P\{X_n = i_n \mid X_{n+1} = i_{n+1}, \cdots, X_{n+m} = i_{n+m}\} = P\{X_n = i_n \mid X_{n+1} = i_{n+1}\}$$

证明 $P\{X_n = i_n \mid X_{n+1} = i_{n+1}, \cdots, X_{n+m} = i_{n+m}\}$

$$= \frac{P\{X_{n+m} = i_{n+m}, X_{n+m-1} = i_{n+m-1}, \cdots, X_{n+1} = i_{n+1}, X_n = i_n\}}{P\{X_{n+m} = i_{n+m}, X_{n+m-1} = i_{n+m-1}, \cdots, X_{n+1} = i_{n+1}\}}$$

$$= \frac{P\{X_{n+m} = i_{n+m} \mid X_{n+m-1} = i_{n+m-1}\} P\{X_{n+m-1} = i_{n+m-1}, \cdots, X_{n+1} = i_{n+1}, X_n = i_n\}}{P\{X_{n+m} = i_{n+m} \mid X_{n+m-1} = i_{n+m-1}\} P\{X_{n+m-1} = i_{n+m-1}, \cdots, X_{n+1} = i_{n+1}\}}$$

$$= \cdots = \frac{P\{X_{n+1} = i_{n+1}, X_n = i_n\}}{P\{X_{n+1} = i_{n+1}\}}$$

$$= P\{X_n = i_n \mid X_{n+1} = i_{n+1}\}$$

性质 2 表明，一个马氏链如果按相反方向的时间排列，所组成的序列也是一个马氏链.

性质 3 设$\{X_n, n \geqslant 0\}$为马氏链，其状态空间为 I，若 $0 \leqslant s \leqslant r < n$，则在 $X_r = i_r$ 的条件下，有

$$P\{X_n = i_n, X_s = i_s \mid X_r = i_r\} = P\{X_n = i_n \mid X_r = i_r\} P\{X_s = i_s \mid X_r = i_r\}$$

证明 $P\{X_n = i_n, X_s = i_s \mid X_r = i_r\} = \dfrac{P\{X_n = i_n, X_s = i_s, X_r = i_r\}}{P\{X_r = i_r\}}$

$$= \frac{P\{X_n = i_n \mid X_s = i_s, X_r = i_r\} P\{X_s = i_s, X_r = i_r\}}{P\{X_r = i_r\}}$$

$$= P\{X_n = i_n \mid X_r = i_r\} P\{X_s = i_s \mid X_r = i_r\}$$

性质 3 表明，若已知“现在”，则“过去”与“将来”是独立的.

性质 4 设$\{X_n, n \geqslant 0\}$为马氏链，其状态空间为 I，则

$$P\{X_{n+1} = i_{n+1}, \cdots, X_{n+m} = i_{n+m} \mid X_n = i_n, \cdots, X_0 = i_0\}$$
$$= P\{X_{n+1} = i_{n+1}, \cdots, X_{n+m} = i_{n+m} \mid X_n = i_n\}$$

证明 $P\{X_{n+1} = i_{n+1}, \cdots, X_{n+m} = i_{n+m} \mid X_n = i_n, \cdots, X_0 = i_0\}$

$$= \frac{P\{X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n, X_{n+1} = i_{n+1}, \cdots, X_{n+m} = i_{n+m}\}}{P\{X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n\}}$$

$$= \frac{p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} p_{i_n i_{n+1}} \cdots p_{i_{n+m-1} i_{n+m}}}{p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}}$$

$$= P\{X_{n+1} = i_{n+1} \mid X_n = i_n\} P\{X_{n+2} = i_{n+2} \mid X_{n+1} = i_{n+1}\}$$
$$\cdots P\{X_{n+m} = i_{n+m} \mid X_{n+m-1} = i_{n+m-1}\}$$

$$= P\{X_{n+1} = i_{n+1}, \cdots, X_{n+m} = i_{n+m} \mid X_n = i_n\}$$

性质 4 表明，若已知“现在”，则“过去”对“将来”各时刻的状态都不产生影响. 特别地，

$$P\{X_{n+m} = i_{n+m} \mid X_n = i_n, \cdots, X_0 = i_0\} = P\{X_{n+m} = i_{n+m} \mid X_n = i_n\}$$

性质 5 设$\{X_n, n \geqslant 0\}$为马氏链，其状态空间为 I，则对任意给定的 n 个整数，$0 \leqslant k_1 < k_2 < \cdots < k_n$，有

$$P\{X_{k_n} = i_{k_n} \mid X_{k_{n-1}} = i_{k_{n-1}}, \cdots, X_{k_1} = i_{k_1}\} = P\{X_{k_n} = i_{k_n} \mid X_{k_{n-1}} = i_{k_{n-1}}\}$$

由性质 5 易知，马氏链的子链也是马氏链.

4.2.5 一些例子

【例 4.3】 (独立随机变量序列和) 设$\{\xi_n, n \geqslant 0\}$为独立同分布随机变量序列，分布律为

$$P\{\xi_n = k\} = q_k, k = 0, 1, \cdots$$

令 $X_n = \sum_{k=0}^{n} \xi_k$，则$\{X_n, n \geqslant 0\}$是一 Markov 链，且

$$p_{ij}=\begin{cases}q_{j-i}, & j\geqslant i\\ 0, & j<i\end{cases}$$

证明：$\{\xi_n,n\geqslant 0\}$是一 Markov 链.

证明 先证$\{X_n,n=0,1,2,\cdots\}$满足 Markov 性. 由$\{\xi_0,\xi_1,\cdots,\xi_n,\cdots\}$相互独立，有

$$\begin{aligned}&P\{X_n=i_n\,|\,X_{n-1}=i_{n-1},\cdots,X_1=i_1,X_0=i_0\}\\&=\frac{P\{X_n=i_n,X_{n-1}=i_{n-1},\cdots,X_1=i_1,X_0=i_0\}}{P\{X_{n-1}=i_{n-1},\cdots,X_1=i_1,X_0=i_0\}}\\&=\frac{P\{\xi_n=i_n-i_{n-1},\xi_{n-1}=i_{n-1}-i_{n-2},\cdots,\xi_1=i_1-i_0,\xi_0=i_0\}}{P(\xi_{n-1}=i_{n-1}-i_{n-2},\cdots,\xi_1=i_1-i_0,\xi_0=i_0)}\\&=P\{\xi_n=i_n-i_{n-1}\}\\&=P\{X_n=i_n\,|\,X_{n-1}=i_{n-1}\}\end{aligned}$$

故$\{X_n,n=0,1,2,\cdots\}$满足 Markov 性，且

$$\begin{aligned}p_{ij}&=P\{X_n=j\,|\,X_{n-1}=i\}=P\{\sum_{k=0}^{n}\xi_k=j\mid\sum_{k=0}^{n-1}\xi_k=i\}\\&=P\{\xi_n=j-i\mid\sum_{k=0}^{n-1}\xi_k=i\}\\&=P\{\xi_n=j-i\}=q_{j-i}\quad(j\geqslant i)\end{aligned}$$

【例 4.4】（天气预报问题）设明天是否有雨仅与今天的天气有关，而与过去的天气无关. 又设今天下雨而明天也下雨的概率为α，而今天无雨明天有雨的概率为β；规定有雨天气为状态 0，无雨天气为状态 1. 因此，问题是两个状态的马尔可夫链. 设$\alpha=0.7,\beta=0.4$，求今天有雨且第 4 天仍有雨的概率.

解 由题设条件，得一步转移概率矩阵为

$$\boldsymbol{P}=\begin{pmatrix}p_{00} & p_{01}\\ p_{10} & p_{11}\end{pmatrix}=\begin{pmatrix}\alpha & 1-\alpha\\ \beta & 1-\beta\end{pmatrix}=\begin{pmatrix}0.7 & 0.3\\ 0.4 & 0.6\end{pmatrix}$$

于是，两步转移概率矩阵为

$$\boldsymbol{P}^{(2)}=\boldsymbol{PP}=\begin{pmatrix}0.7 & 0.3\\ 0.4 & 0.6\end{pmatrix}\begin{pmatrix}0.7 & 0.3\\ 0.4 & 0.6\end{pmatrix}=\begin{pmatrix}0.61 & 0.39\\ 0.52 & 0.48\end{pmatrix}$$

四步转移概率矩阵为

$$\boldsymbol{P}^{(4)}=\boldsymbol{P}^{(2)}\boldsymbol{P}^{(2)}=\begin{pmatrix}0.5749 & 0.4251\\ 0.5668 & 0.4332\end{pmatrix}$$

从而得到今天有雨且第 4 天仍有雨的概率为$p_{00}^{(4)}=0.5749$.

【例 4.5】（竞赛问题）甲乙两人进行一种比赛，设每局比赛甲胜的概率是p，乙胜的概率是q，和局的概率为r，且$p+q+r=1$. 设每局比赛胜者记 1 分，负者记 -1 分，和局记零分. 当有一人获得 2 分时，比赛结束. 以X_n表示比赛至n局时甲获得的分数，则$\{X_n,n\geqslant 1\}$是齐次马尔可夫链.

(1) 写出状态空间 I；

(2) 求出两步转移概率矩阵；

(3) 求甲已获 1 分时，再赛两局可以结束比赛的概率.

解 (1) $I=\{-2,-1,0,1,2\}$.

(2) 显然，一步转移概率矩阵为

$$
\boldsymbol{P}=\begin{pmatrix}1&0&0&0&0\\q&r&p&0&0\\0&q&r&p&0\\0&0&q&r&p\\0&0&0&0&1\end{pmatrix}
$$

于是，两步转移概率矩阵为

$$
\boldsymbol{P}^{(2)}=\boldsymbol{PP}
$$

$$
=\begin{pmatrix}1&0&0&0&0\\q+rp&r^2+pq&2pr&p^2&0\\q^2&2rq&r^2+2pq&2pr&0\\0&q^2&2qr&pq+r^2&p+pr\\0&0&0&0&1\end{pmatrix}
$$

(3) 经 2 局结束比赛包括两种情形：甲得 1 分经两步转移至得 2 分而结束比赛，或甲得 1 分经两步转移至 -2 分(乙得 2 分)而结束比赛. 因此，有

$$
p=p_{45}^{(2)}+p_{41}^{(2)}=(p+pr)+0=p(1+r).
$$

在解此题时，常常会漏掉第二种情形. 尽管本题中第二种情形的概率为零，但必须学会全面、准确地分析问题.

【例 4.6】 (广告效益的推算) A 种啤酒的广告改变广告方式后，经市场调查发现：买 A 种啤酒及另三种啤酒 B,C,D (设市场上只有这四种啤酒) 的顾客每两个月的平均转移率如下：

$A\to A(95\%)\to B(2\%)\to C(2\%)\to D(1\%)$,

$B\to A(30\%)\to B(60\%)\to C(6\%)\to D(4\%)$,

$C\to A(20\%)\to B(10\%)\to C(70\%)\to D(0\%)$,

$D\to A(20\%)\to B(20\%)\to C(10\%)\to D(50\%)$,

设目前购买 A,B,C,D 四种啤酒的顾客的分布为 $(25\%,30\%,35\%,10\%)$，求半年后 A 种啤酒占有的市场份额.

解 显然一步转移概率矩阵为

$$
\boldsymbol{P}=\begin{pmatrix}0.95&0.02&0.02&0.01\\0.30&0.60&0.06&0.04\\0.20&0.10&0.70&0.00\\0.20&0.20&0.10&0.50\end{pmatrix}
$$

再令 $\boldsymbol{\mu}=(\mu_1,\mu_2,\mu_3,\mu_4)=(0.25,0.30,0.35,0.10)$.

半年以后顾客的转移概率矩阵为 $\boldsymbol{P}^{(3)}$，而

$$\boldsymbol{P}^{(3)}=\boldsymbol{P}^3=\begin{pmatrix}0.88937 & 0.04587 & 0.04656 & 0.01820\\ 0.60175 & 0.25590 & 0.09880 & 0.04355\\ 0.48340 & 0.13880 & 0.36580 & 0.01196\\ 0.50090 & 0.21340 & 0.14264 & 0.14306\end{pmatrix}$$

因为只关心 A,B,C,D 四种啤酒经三次转移后转到 A 种的概率，所以由 $\boldsymbol{P}^{(3)}$ 的第一列，得

$$(0.25,0.30,0.35,0.10)\begin{pmatrix}0.88937\\ 0.60175\\ 0.48340\\ 0.50090\end{pmatrix}\approx 0.622$$

所以 A 种啤酒在半年后占有的市场份额为62.2%，广告的效益很好.

【例4.7】 （随机游动问题）随机游动是马尔可夫链的经典例子，它的统计特性由它的边界行为所决定，很多实际应用例子都可以看作是随机游动的模型.常见的随机游动有以下几种类型：

(1) 自由随机游动

设有一质点在数轴上随机游动，每隔一单位时间 Δt（设 $\Delta t=1$）移动一次，每次只能向左或向右移动 Δx 单位（设 $\Delta x=1$）.设质点在0时刻的位置为 b，它向左移动的概率为 $p\geqslant 0$，向右移动的概率为 $q\geqslant 0$ $(p+q=1)$，且各次移动相互独立.以 $X(n)$ 表示时刻 n 质点所处的位置，则 $\{X(n),n=0,1,2,\cdots\}$ 是一齐次马氏链，其状态空间为 $I=\{\cdots,-2,-1,0,1,2,\cdots\}$，一步转移概率为

$$\begin{cases}p_{i,i-1}=p, & i\in I,0<p<1\\ p_{i,i+1}=q=1-p, & i\in I,0<p<1\\ p_{ij}=0, & i\neq i+1,i-1,j\in I\end{cases}$$

其一步转移概率矩阵为

$$\boldsymbol{P}=\begin{pmatrix}\ddots & \ddots & \ddots & & & & \\ & p & 0 & q & & & \\ & & p & 0 & q & & \\ & & & p & 0 & q & \\ & & & & \ddots & \ddots & \ddots\end{pmatrix}$$

(2) 带有一个吸收壁的随机游动

设(1)中的随机游动限制在 $I=\{0,1,2,\cdots\}$ 内，当质点移动到状态0后就永远停留在该位置，即 $p_{00}=1$，其余 $p_{ij}(i,j\geqslant 1)$ 同(1).这时序列 $\{X(n),n\geqslant 0\}$ 称为带有一个吸收壁0的随机游动，其特点是：当 $X(n)=0$ 时，$X(n+1)$ 就停留在零状态.

此时$\{X(n),n=0,1,2,\cdots\}$是一齐次马氏链,其状态空间为$I=\{0,1,2,\cdots\}$,一步转移概率为

$$\begin{cases} p_{i,i-1}=p, & i\geqslant 1,i\in I \\ p_{i,i+1}=q, & i\geqslant 1,i\in I \\ p_{ij}=0, & j\neq i+1,i-1,i\geqslant 1,i\in I \\ p_{00}=1 \end{cases}$$

其一步转移概率矩阵为

$$\boldsymbol{P}=\begin{pmatrix} 1 & 0 & & & \\ p & 0 & q & & \\ & p & 0 & q & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

(3) 带有两个吸收壁的随机游动

设(1)中的随机游动限制在$I=\{0,1,2,\cdots,m\}$内,当质点移动到状态0或m后就永远停留在该位置,即$p_{00}=1,p_{mm}=1$,其余$p_{ij}(1\leqslant i,j\leqslant m-1)$同(1).这时序列$\{X(n),n\geqslant 0\}$称为带有两个吸收壁0和$m$的随机游动.此时$\{X(n),n=0,1,2,\cdots\}$是一齐次马氏链,状态空间为$I=\{0,1,2,\cdots,m\}$,0,$m$为两个吸收状态,它的一步转移概率为

$$\begin{cases} p_{i,i-1}=p \quad (1\leqslant i\leqslant m-1) \\ p_{i,i+1}=q=1-p \quad (1\leqslant i\leqslant m-1) \\ p_{ij}=0 \quad (j\neq i+1,i-1;1\leqslant i\leqslant m-1) \\ p_{00}=1 \\ p_{mm}=1 \\ p_{0j}=0 \quad (j\neq 0) \\ p_{mj}=0 \quad (j\neq m) \end{cases}$$

其一步转移概率矩阵为

$$\boldsymbol{P}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ p & 0 & q & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & p & 0 & q & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & p & 0 & q \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{pmatrix}_{(m+1)\times(m+1)}$$

(4) 带有一个反射壁的随机游动

若状态空间为$I=\{0,1,2,\cdots\}$,且$P_{00}=0,P_{01}=1$,则序列$\{X(n),n\geqslant 0\}$称为带有一个反射壁(状态0)的随机游动.它的一步转移概率为

$$\begin{cases} p_{i,i-1} = p & (i \geqslant 1) \\ p_{i,i+1} = q = 1-p & (i \geqslant 1) \\ p_{ij} = 0 & (j \neq i+1, i-1; i \geqslant 1) \\ p_{01} = 1 \\ p_{00} = 0 \end{cases}$$

其一步转移概率矩阵为

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 & & & \\ p & 0 & q & & \\ & p & 0 & q & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

(5) 带有两个反射壁的随机游动

此时的状态空间为 $I=\{0,1,2,\cdots,m\}$，且 $p_{00}=0, p_{01}=1, p_{mm}=0, p_{m,m-1}=1$，两个反射壁为状态 0 和 m. 则其一步转移概率为

$$\begin{cases} p_{i,i-1} = p & (1 \leqslant i \leqslant m-1) \\ p_{i,i+1} = q = 1-p & (1 \leqslant i \leqslant m-1) \\ p_{ij} = 0 & (j \neq i+1, i-1; 1 \leqslant i \leqslant m-1) \\ p_{00} = 0 \\ p_{01} = 1 \\ p_{mm} = 0 \\ p_{m,m-1} = 1 \end{cases}$$

则它的一步转移概率矩阵为

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ p & 0 & q & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & p & 0 & q & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & p & 0 & q \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{pmatrix}_{(m+1)\times(m+1)}$$

(6) 带有一个弹性壁的随机游动

此时的状态空间为 $I=\{0,1,2,\cdots\}$，且 $p_{00}=\alpha, p_{01}=\beta(\alpha+\beta=1)$，$0<\alpha<1$，一个弹性壁为状态 0. 则其一步转移概率为

$$\begin{cases} p_{i,i-1} = p \\ p_{i,i+1} = q = 1-p \\ p_{00} = \alpha \\ p_{01} = \beta, i \geqslant 1 \\ p_{ij} = 0 \end{cases}$$

其一步转移概率矩阵为

$$P = \begin{pmatrix} \alpha & \beta & & & \\ p & 0 & q & & \\ & p & 0 & q & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

【例 4.8】 ($M/G/1$ 排队系统)假设顾客依参数为 λ 的 Poisson 过程来到只有一个服务员的服务站,若服务员空闲来客就立刻得到服务,否则排队等待直至轮到他.设每名顾客接受服务的时间独立同分布,分布函数为 $G(x)$,且与顾客到达过程相互独立.这个系统称为 $M/G/1$ 排队系统,字母 M 表示顾客到达的时间间隔服从指数分布,G 表示服务时间的分布,1 表示单个服务员.

令 X_n 表示第 n 个顾客服务完毕时等待接受服务的顾客数,再令 U_n 为第 n 个顾客接受服务的时间内来到服务机构的顾客数,则

$$X_{n+1} = X_n - \delta(X_n) + U_{n+1}$$

其中 $\delta(x) = \begin{cases} 1, x \neq 0 \\ 0, x = 0 \end{cases}$.

则 $\{X_n, n = 0,1,2,\cdots\}$ 是齐次 Markov 链,其一步转移概率为

$$p_{ij}(n) = P\{U_{n+1} = j - i + \delta(i)\}, \forall n \geqslant 1$$

证明 由于 $X_{n+1} = X_n - \delta(X_n) + U_{n+1}$,且 U_{n+1} 与 $\{X_k, k = 1,2,\cdots,n\}$ 相互独立,从而

$$\begin{aligned} & P\{X_{n+1} = i_{n+1} \mid X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_1 = i_1\} \\ & = \frac{P\{X_n - \delta(X_n) + U_{n+1} = i_{n+1}, X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_1 = i_1\}}{P\{X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_1 = i_1\}} \\ & = \frac{P\{U_{n+1} = i_{n+1} - i_n + \delta(i_n), X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_1 = i_1\}}{P\{X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_1 = i_1\}} \\ & = P\{U_{n+1} = i_{n+1} - i_n + \delta(i_n)\}, \forall n \geqslant 1 \end{aligned}$$

同理可证

$$P\{X_{n+1} = i_{n+1} \mid X_n = i_n\} = P\{U_{n+1} = i_{n+1} - i_n + \delta(i_n)\}, \forall n \geqslant 1$$

从而 $\{X_n, n = 0,1,2,\cdots\}$ 满足 Markov 性.再注意到顾客接受服务的时间是独立同分布的,从而其一步转移概率与 n 无关,即该 Markov 链也是齐次的.

【例 4.9】 (离散分枝过程)离散分枝过程是 Markov 过程的重要特例,常用来描述细胞分裂,种群繁衍,粒子裂变等现象,在随机过程的理论和应用中占有非常重要的地位.下面介绍的模型是英国博物学家 Galton Watson 在研究家族谱系关系时引入的,因此也称为 Galton-Watson 分枝过程(简称 G-W 过程).

考虑一个能产生同类后代的个体组成的群体.每个个体以概率 $p_m, m \geqslant 0$ 产生 m 个新后代,且与别的个体产生的后代个数相互独立.初始的个体个数为 X_0,其后代构成第一代,总数记为 X_1.以此类推,以 X_n 表示第 n 代的总数,记 $\xi_i^{(n)}$ 表示第 n 代

的第 i 个个体产生的后代个数，那么 $\{X_n, n=0,1,2,\cdots\}$ 就为一个 Markov 链，称之为时间离散的分枝过程.

设 $\{\xi_i^{(n)}, i=1,2,\cdots; n=1,2,\cdots\}$ 为概率空间 (Ω,Σ,P) 上的一族独立同分布的取非负整数的随机变量，其公共分布为 $\{a_k = P\{\xi_i^{(n)}=k\}, k\geqslant 0\}$. 若定义 X_0 为 $\mathbf{R}$ 上的任意取正整数的随机变量且

$$X_n = \sum_{i=1}^{X_{n-1}} \xi_i^{(n)}, n=1,2,\cdots$$

则称 $\{X_n, n=0,1,2,\cdots\}$ 是一个离散时间的分枝过程，或分枝链. 则 $\{X_n, n=0,1,2,\cdots\}$ 为齐次 Markov 链，且

$$p_{ij} = \left.\frac{\partial^j\left[\left(\sum_{k=0}^{+\infty} a_k s^k\right)^i\right]}{j!\partial s^j}\right|_{s=0}$$

证明　先证 $\{X_n, n=0,1,2,\cdots\}$ 为齐次 Markov 链.

注意到 X_n 只依赖于独立同分布的随机变量 $\{\xi_i^{(n)}, i=1,2,\cdots; n=1,2,\cdots\}$，从而

$$\begin{aligned}
&P\{X_n = i_n \mid X_{n-1}=i_{n-1},\cdots,X_1=i_1,X_0=i_0\} \\
&= \frac{P\{X_n=i_n, X_{n-1}=i_{n-1},\cdots,X_1=i_1,X_0=i_0\}}{P\{X_{n-1}=i_{n-1},\cdots,X_1=i_1,X_0=i_0\}} \\
&= \frac{P\{\sum_{k=1}^{i_{n-1}} \xi_k^{(n)} = i_n, X_{n-1}=i_{n-1},\cdots,X_1=i_1,X_0=i_0\}}{P\{X_{n-1}=i_{n-1},\cdots,X_1=i_1,X_0=i_0\}} \\
&= P\{\sum_{k=1}^{i_{n-1}} \xi_k^{(n)} = i_n\} = P\{X_n=i_n \mid X_{n-1}=i_{n-1}\}
\end{aligned}$$

因为 $\{\xi_i^{(n)}, i=1,2,\cdots; n=1,2,\cdots\}$ 独立同分布，从而转移概率 p_{ij} 与 n 无关，即 $\{X_n, n=0,1,2,\cdots\}$ 为齐次 Markov 链.

由母函数的性质（独立同分布随机变量和的母函数等于母函数之积）知，$\sum_{j=0}^{+\infty} p_{ij}s^j = [f(s)]^i$，其中 $f(s)$ 为分布为 $\{a_k = P\{\xi_i^{(n)}=k\}, k\geqslant 0\}$ 的母函数，即 $f(s) = \sum_{k=0}^{+\infty} a_k s^k$. 进一步，再由母函数与概率分布的关系得

$$p_{ij} = \left.\frac{\partial^j\left(\sum_{k=0}^{+\infty} p_{ik} s^k\right)}{j!\partial s^j}\right|_{s=0} = \left.\frac{\partial^j\left[\left(\sum_{k=0}^{+\infty} a_k s^k\right)^i\right]}{j!\partial s^j}\right|_{s=0}$$

4.3 Markov 链的状态分类及性质

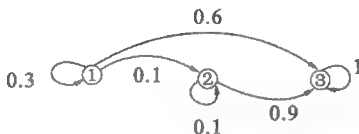
4.3.1 状态分类

本节只讨论齐次马尔可夫链$\{X_n,n\geqslant 0\}$的状态分类,设状态空间为$I=\{0,1,2,\cdots\}$,一步转移概率为$p_{ij},i,j\in I$.

下面由一个简单的例子引入.假设系统有三种可能状态$I=\{1,2,3\}$."1"表示系统运行良好,"2"表示系统运行不正常,"3"表示系统失效.以X_n表示系统在时刻n的状态,并设$\{X_n,n\geqslant 0\}$是一Markov链.其一步转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix}0.3 & 0.1 & 0.6\\ 0 & 0.1 & 0.9\\ 0 & 0 & 1\end{bmatrix}$$

用状态转移图表示为



可见,从状态"1"或"2"出发经有限次转移后总要到达状态"3",而一旦到达状态"3"则永远停在状态"3".显然状态"1","2"与状态"3"概率性质不同.因此,可以依概率性质对状态进行分类.

定义 4.5 设$\{X_n,n\geqslant 0\}$是Markov链,$i,j\in I$,若存在$n\geqslant 0$使得$p_{ij}^{(n)}>0$,则称i可达j,记为$i\rightarrow j$;若$i\rightarrow j$,且$j\rightarrow i$,则称i和j是互通的,记为$i\leftrightarrow j$.

定理 4.3 互通是一种等价关系,即满足

(1) 自反性:$i\leftrightarrow i$.

(2) 对称性:若$i\leftrightarrow j$,则$j\leftrightarrow i$.

(3) 传递性:若$i\leftrightarrow j,j\leftrightarrow k$,则$i\leftrightarrow k$.

证明 如果$i\rightarrow k,k\rightarrow j$,则由定义知,存在$r\geqslant 1$和$n\geqslant 1$,使得

$$p_{ik}^{(r)}>0,p_{kj}^{(n)}>0$$

根据C-K方程,有

$$p_{ij}^{(r+n)}=\sum_{m\in I}p_{im}^{(r)}p_{mj}^{(n)}\geqslant p_{ik}^{(r)}p_{kj}^{(n)}>0(k\in I)$$

因此,$i\rightarrow j$.同理可以证明互通的情形.

注：根据等价关系的性质，可以把两个互通的状态归为一类，则同一类的状态是互通的，并且任何一个状态不可能同时属于两个不同的类.

【例 4.10】　设 $I=\{0,1,2\}$，一步转移概率矩阵为

$$P=\begin{pmatrix}1/2 & 1/2 & 0\\ 1/2 & 1/4 & 1/4\\ 0 & 1/3 & 2/3\end{pmatrix}$$

画出状态传递图并分析各状态间的关系.

解　状态传递图如图 4.1 所示.

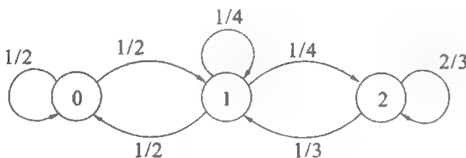


图 4.1

由状态 0 可达状态 2，$⓪\xrightarrow{1/2}①\xrightarrow{1/4}②$；

由状态 2 可达状态 0，$②\xrightarrow{1/3}①\xrightarrow{1/2}⓪$.

定义 4.6　若 Markov 链只存在一个类，则称它是不可约的马氏链. 否则称 Markov 链是可约的.

定义 4.7　设 $C\subset I$ 是状态空间的一个子集，若对 $\forall i\in C,\forall j\notin C$，$i$ 不可达 j，则称 C 是 I 的一个闭集.

注：若 C 是一个闭集，则对任意 n，$\sum_{j\in C}p_{ij}^{(n)}=1$.

【例 4.11】　请分析下面 Markov 链的状态的类.

(1) $P=\begin{pmatrix}1/2 & 1/2 & 0\\ 1/2 & 1/4 & 1/4\\ 0 & 2/3 & 1/3\end{pmatrix}$；

(2) $P=\begin{pmatrix}0.5 & 0.5 & 0 & 0\\ 0.5 & 0.5 & 0 & 0\\ 0.25 & 0.25 & 0.25 & 0.25\\ 0 & 0 & 0 & 1\end{pmatrix}$.

解　(1) 是不可约 Markov 链；

(2) 有 3 个类，$\{0,1\}$，$\{2\}$，$\{3\}$，其中 $\{0,1\}$ 是一个闭集.

定义 4.8　对于任意的 $i,j\in I$，称：

$$T_{ij}\triangleq\min\{n:X_0=i,X_n=j,n\geqslant 1\}$$

为从状态 i 出发首次到达(进入)状态 j 的时间(时刻)，简称首达时间.

首达时间 T_{ij} 是一随机变量,它取值于 $N_{+\infty}=\{1,2,\cdots,+\infty\}$.

定义 4.9 对于任意的 $i,j\in I$,称

$$f_{ij}^{(n)}\triangleq P\{T_{ij}=n|X_0=i\}$$

为系统在 0 时从状态 i 出发,经 n 步首次到达状态 j 的概率.

由定义 4.9,显然有

$$f_{ij}^{(n)}=P\{X_n=j;X_m\neq j,m=1,2,\cdots,n-1|X_0=i\}$$

$$f_{ij}^{(1)}=p_{ij}=P\{X_1=j|X_0=i\}$$

$$f_{ij}^{(+\infty)}=P\{X_m\neq j,\forall m\geqslant 1|X_0=i\}$$

定义 4.10 对于任意的 $i,j\in I$,称

$$f_{ij}=\sum_{1\leqslant n<+\infty}f_{ij}^{(n)}=\sum_{1\leqslant n<+\infty}P\{T_{ij}=n|X_0=i\}=P\{T_{ij}<+\infty\}$$

为系统在 0 时从状态 i 出发,经过有限步转移后迟早到达状态 j 的概率.

$P\{T_{ij}=+\infty\}\triangleq f_{ij}^{(+\infty)}=1-f_{ij}$,它表示系统在0时从状态$i$出发,经过有限步转移后不可能到达状态 j 的概率.

下面给出首达概率如下的基本性质:

定理 4.4 (1) 首达概率可以用一步转移概率表示,即

$$f_{ij}^{(n)}=\sum_{i_1\neq j}\sum_{i_2\neq j}\cdots\sum_{i_k\neq j}p_{ii_1}p_{i_1i_2}\cdots p_{i_{n-1}j}$$

(2) 对于任意的 $i,j\in I$,首达概率与 n 步转移概率之间有如下关系:

$$p_{ij}^{(n)}=\sum_{l=1}^{n}f_{ij}^{(l)}p_{jj}^{(n-l)}=\sum_{l=1}^{n}f_{ij}^{(n-l)}p_{jj}^{(l)},1\leqslant n<+\infty$$

(3) 对于任意的 $i,j\in I$,有 $0\leqslant f_{ij}^{(n)}<f_{ij}\leqslant 1$,当 $i=j$ 时,T_{ii} 是从 i 出发,首次返回 i 的时刻;f_{ii} 是从状态 i 出发,迟早要返回状态 i 的概率.

(4) 对于任意的 $i,j\in I,f_{ij}>0\Leftrightarrow i\rightarrow j$.

(5) 对于任意的 $i,j\in I,f_{ij}>0$ 且 $f_{ji}>0\Leftrightarrow i\leftrightarrow j$.

证明 结论(1)与(3)显然成立.下面只证结论(2)和(4),(5)是结论(4)的推广.考虑先证明(2):

$$\begin{aligned}P\{X_n=j\mid X_0=i\}&=P\{\bigcup_{l=1}^{n}T_{ij}=l,X_n=j\mid X_0=i\}\\&=\sum_{l=1}^{n}P\{T_{ij}=l,X_n=j\mid X_0=i\}\\&=\sum_{l=1}^{n}\frac{P\{T_{ij}=l,X_n=j,X_0=i\}}{P\{X_0=i\}}\\&=\sum_{l=1}^{n}P\{X_n=j\mid T_{ij}=l,X_0=i\}P\{T_{ij}=l\mid X_0=i\}\end{aligned}$$

$$= P\{X_n = j \mid X_l = j, X_k \neq j, k = 1,2,\cdots,l-1, X_0 = i\} f_{ij}^{(l)}$$
$$= \sum_{l=1}^{n} f_{ij}^{(l)} P\{X_n = j \mid X_l = j\}$$
$$= \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$$

再证明(4)：当 $i \to j$ 时，$\exists n > 0$，使得 $p_{ij}^{(n)} > 0$，取 $n' = \min\{n: p_{ij}^{(n)} > 0\}$，则有

$$f_{ij}^{(n')} = P\{T_{ij} = n' \mid X_0 = i\} = p_{ij}^{(n')} > 0$$

因此

$$f_{ij} = \sum_{n=1}^{+\infty} f_{ij}^{(n)} \geqslant f_{ij}^{(n')} > 0$$

反之，当 $f_{ij} > 0$ 时，$\exists n' > 0$，使得 $f_{ij}^{(n')} > 0$，从而 $p_{ij}^{(n')} > 0$，得 $i \to j$.

4.3.2　常返性判别及其性质

定义 4.11　（周期性）若集合 $\{n: n \geqslant 1, p_{ii}^{(n)} > 0\}$ 非空，则称它的最大公约数 $d = d(i)$ 为状态 i 的周期. 若 $d > 1$，则 i 是周期的；若 $d = 1$，则称 i 是非周期的. 并特别规定，若上述集合为空集，则称 i 的周期为无穷大.

定理 4.5　周期性是类的性质，即若状态 i 和 j 同属一类，则 $d(i) = d(j)$.

证明　由于 $i \leftrightarrow j$，即存在 m、n 使得 $p_{ij}^{(m)} > 0$ 且 $p_{ji}^{(n)} > 0$，因此由 C-K 方程有

$$p_{ii}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{ki}^{(n)} > 0$$

对所有满足 $p_{jj}^{(s)} > 0$ 的 s 都有

$$p_{ii}^{(n+s+m)} \geqslant p_{ij}^{(n)} p_{jj}^{(s)} p_{ji}^{(m)} > 0$$

因此 $m + n$ 和 $m + n + s$ 都是 $d(i)$ 的整数倍，故 $d(i)$ 可以整除 s. 而 s 是 $d(j)$ 的整数倍，所以 $d(i)$ 也整除 $d(j)$. 反过来也可以证明 $d(j)$ 整除 $d(i)$. 于是 $d(i) = d(j)$.

定义 4.12　对于状态 $i \in I$，如果 $f_{ii} = 1$，则称状态 i 为**常返态（返回态）**；如果 $f_{ii} < 1$，则称状态 i 为**非常返态（滑过态）**.

注：对于任何状态 i, j，以 $f_{ij}^{(n)}$ 表示从 i 出发经 n 步后首次到达 j 的概率，则有

$$f_{ij}^{(0)} = 0$$
$$f_{ij}^{(n)} = P\{X_n = j, X_k \neq j, \forall 0 < k < n \mid X_0 = i\},\ n \geqslant 1$$

容易看出，集合 $A_n = \{X_n = j, X_k \neq j, \forall 0 < k < n \mid X_0 = i\}$ 在 n 不同时不相交，并且 $\bigcup_{n=1}^{+\infty} A_n$ 表示总有一个 n 使得过程经过 n 步后可从 i 到达 j. $f_{ij} = \sum_{n=0}^{+\infty} f_{ij}^{(n)}$ 表示从 i 出发，在有限步内可以到达 j 的概率. 因此，易得

$$i \text{ 是常返状态} \Leftrightarrow f_{ii} = 1$$

$$i\text{ 是非常返状态}\Leftrightarrow f_{ii}<1$$

令 $\mu_{ij}\triangleq E\{T_{ij}\mid X_0=i\}=\sum\limits_{n=1}^{+\infty}nf_{ij}^{(n)}$，其中 μ_{ij} 表示从状态i出发，首次到达状态j的平均转移步数(时间)．特别地，若 $i=j$，则 $\mu_{ii}\triangleq\mu_i$ 是从状态i出发，首次返回状态 i 的平均转移步数，称为状态 i 的**平均返回时间**；对应的 f_{ii} 称为状态i 的**返回概率**；$f_{ii}^{(n)}$ 称为从状态i 出发，经过 n 步转移首次返回状态 i 的概率．

对于常返态 $i\in I$，若 $\mu_i<+\infty$，则称i是**正常返**的．否则称i是**零常返**的．特别地，若 i 是正常返且周期为 1，则称 i 是**遍历的**．若 i 是遍历的，且 $f_{ii}^{(1)}=1$，则称 i 是**吸收状态**．

定理 4.6 i 是常返状态 $\Leftrightarrow f_{ii}=1\Leftrightarrow\sum\limits_{n=0}^{+\infty}p_{ii}^{(n)}=+\infty$；$i$ 是非常返状态 $\Leftrightarrow f_{ii}<1\Leftrightarrow\sum\limits_{n=0}^{+\infty}p_{ii}^{(n)}=\dfrac{1}{1-f_{ii}}<+\infty$，进而有 $\lim\limits_{n\to+\infty}p_{ii}^{(n)}=0$．

证明 首先证明，对任意状态 i,j 及 $1\leqslant n<+\infty$，有

$$p_{ij}^{(n)}=\sum_{k=1}^{n}f_{ij}^{(k)}p_{jj}^{(n-k)}$$

事实上，由 C-K 方程可得

$$\begin{aligned}
p_{ij}^{(n)}&=P\{X_n=j\mid X_0=i\}\\
&=\sum_{k=1}^{n}P\{\{X_n\}\text{ 在时刻 }k\text{ 首次到达 }j\text{，再从时刻 }k\text{ 经 }n-k\text{ 步后到达 }j\}\\
&=\sum_{k=1}^{n}P\{X_n=j\mid X_k=j,X_l\neq j,0<l<k,X_0=i\}\\
&\quad\cdot P\{X_k=j,X_l\neq j,0<l<k\mid X_0=i\}\\
&=\sum_{k=1}^{n}P\{X_n=j\mid X_k=j\}P\{X_k=j,X_l\neq j,0<l<k\mid X_0=i\}\\
&=\sum_{k=1}^{n}f_{ij}^{(k)}p_{jj}^{(n-k)}
\end{aligned}$$

因此，

$$\begin{aligned}
\sum_{n=0}^{+\infty}p_{ii}^{(n)}&=p_{ii}^{(0)}+\sum_{n=1}^{+\infty}\sum_{k=1}^{n}f_{ii}^{(k)}p_{ii}^{(n-k)}\\
&=1+\sum_{k=1}^{+\infty}\sum_{n=k}^{+\infty}f_{ii}^{(k)}p_{ii}^{(n-k)}\\
&=1+\sum_{k=1}^{+\infty}f_{ii}^{(k)}\sum_{n=k}^{+\infty}p_{ii}^{(n-k)}\\
&=1+f_{ii}\sum_{n=0}^{+\infty}p_{ii}^{(n)}
\end{aligned}$$

所以，$\sum_{n=0}^{+\infty} p_{ii}^{(n)} = \frac{1}{1-f_{ii}}$，从而得到如下结论：

(1) $\sum_{n=0}^{+\infty} p_{ii}^{(n)} = \frac{1}{1-f_{ii}} < +\infty \Leftrightarrow f_{ii} < 1 \Leftrightarrow i$ 是非常返状态，且考虑级数 $\sum_{n=0}^{+\infty} p_{ii}^{(n)}$ 是收敛的，因而有 $\lim_{n\to+\infty} p_{ii}^{(n)} = 0$；

(2) $\sum_{n=0}^{+\infty} p_{ii}^{(n)} = +\infty \Leftrightarrow f_{ii} = 1 \Leftrightarrow i$ 是常返状态.

定理 4.7　常返性是一个类的性质，即若状态 i 和 j 同属一类，则 i 和 j 同为常返或非常返态.

证明　由于 $i \leftrightarrow j$，故存在 n 和 m，使得 $p_{ij}^{(n)} > 0, p_{ji}^{(m)} > 0$，由 C-K 方程有

$$p_{ii}^{(n+m+l)} \geqslant p_{ij}^{(n)} p_{jj}^{(l)} p_{ji}^{(m)}$$
$$p_{jj}^{(n+m+l)} \geqslant p_{ji}^{(m)} p_{ii}^{(l)} p_{ij}^{(n)}$$

求和得

$$\sum_{l=0}^{+\infty} p_{ii}^{(n+m+l)} \geqslant p_{ij}^{(n)} p_{ji}^{(m)} \sum_{l=0}^{+\infty} p_{jj}^{(l)}$$
$$\sum_{l=0}^{+\infty} p_{jj}^{(n+m+l)} \geqslant p_{ji}^{(m)} p_{ij}^{(n)} \sum_{l=0}^{+\infty} p_{ii}^{(l)}$$

因此，$\sum_{l=0}^{+\infty} p_{jj}^{(l)}$ 和 $\sum_{l=0}^{+\infty} p_{ii}^{(l)}$ 是相互控制的，同为无穷或有限，故 i 和 j 同为常返或非常返态.

归纳状态分类图如图 4.2 所示.

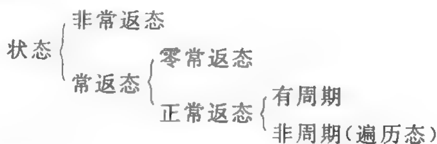


图 4.2

【例 4.12】　一个状态空间为 $I = \{0,1,2,3\}$ 的马尔可夫链，其一步转移概率矩阵为

$$\boldsymbol{P} = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

试对其状态进行分类.

解　这是一个有限状态的马尔可夫链，所有状态都是相通的，因此所有状态都是常返态. 其状态传递图如图 4.3 所示.

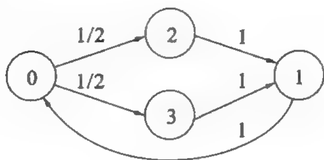


图 4.3

【例 4.13】 设 $I=\{1,2,3,4\}$，其一步转移概率矩阵为

$$P=\begin{pmatrix}1/2 & 1/2 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 1/3 & 2/3 & 0\\ 1/2 & 0 & 1/2 & 0\end{pmatrix}$$

试对其状态进行分类，确定哪些状态是常返态，并确定其周期.

解 画状态传递图如图 4.4 所示，

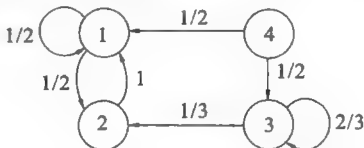


图 4.4

因为对一切 $n\geqslant 1$，$f_{44}^{(n)}=0$，所以 $f_{44}=0<1$. 从而知状态 4 是非常返态. 又 $f_{33}^{(1)}=\dfrac{2}{3}$，$f_{33}^{(n)}=0(n\geqslant 2)$，所以 $f_{33}=\dfrac{2}{3}<1$. 从而知状态 3 也是非常返态. 而

$$f_{11}=f_{11}^{(1)}+f_{11}^{(2)}=1/2+1/2=1$$

$$f_{22}=f_{22}^{(1)}+f_{22}^{(2)}+\cdots=0+1/2+1/2^2+\cdots=1$$

所以状态 1 和状态 2 是常返态. 又知

$$\mu_1=\sum_{n=1}^{+\infty}nf_{11}^{(n)}=1\cdot\frac{1}{2}+2\cdot\frac{1}{2}=\frac{3}{2}<+\infty$$

$$\mu_2=\sum_{n=1}^{+\infty}nf_{22}^{(n)}=1\cdot 0+2\cdot\frac{1}{2}+3\cdot\frac{1}{2^2}+\cdots=3<+\infty$$

而其周期均为 1，故状态 1 与状态 2 是正常返态，且为遍历态.

【例 4.14】 设马尔可夫链的状态空间 $I=\{0,1,\cdots\}$，其转移概率为 $p_{i,i+1}=1/2$，$p_{i0}=1/2$，$i=0,1,2,\cdots$，试画出状态传递图并对其状态进行分类.

解 状态传递图如图 4.5 所示.

$$f_{00}=\sum_{n=0}^{+\infty}f_{00}^{(n)}=\sum_{n=0}^{+\infty}\frac{1}{2^n}=1$$

故状态 0 是常返态.

又任一状态 i 都与 0 相通，所以由判别准则知，任一状态 i 都是常返态. 而 $\mu_0=$

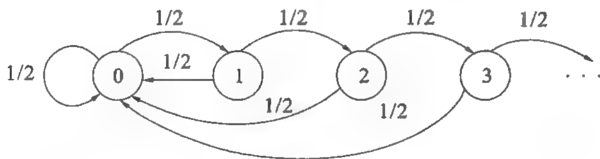


图 4.5

$\sum_{n=1}^{+\infty} n \cdot \frac{1}{2^n} < +\infty$,所以状态0是常返态;又 $p_{00} = \frac{1}{2} > 0$,故状态0是非周期的,而且是遍历的.故依判别准则,任一状态 i 都是遍历的.

当状态较多时,逐个对状态进行计算是困难的.如果状态是相通的,选择其中一个易于识别的状态进行计算和识别,也就识别了其余的状态,从而可以大大减少工作量.

【例 4.15】 已知马尔可夫链 $\{X_n, n \geqslant 0\}$ 的状态空间 $I = \{1,2,3,4\}$,其转移概率矩阵 $\boldsymbol{P}$ 的状态转移图如图 4.6 所示,试对其状态进行分类.

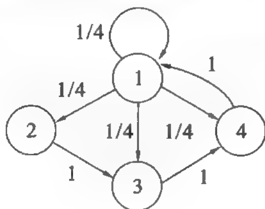


图 4.6

解 因为 $p_{11} = 1/4$,故 $\{n, p_{11}^{(n)} > 0\}$ 的最大公约数 $d = 1$,所以状态1是非周期的.从而每一状态都是非周期的.

考虑状态1是否是常返态.因为 $f_{11}^{(1)} = 1/4$,而

$$
\begin{aligned}
f_{11}^{(2)} &= P\{X_2 = 1, X_1 \neq 1 \mid X_0 = 1\} \\
&= P\{X_2 = 1, X_1 = 2 \mid X_0 = 1\} + P\{X_2 = 1, X_1 = 3 \mid X_0 = 1\} + \\
&\quad\ P\{X_2 = 1, X_1 = 4 \mid X_0 = 1\} \\
&= P\{X_2 = 1 \mid X_1 = 4, X_0 = 1\} P\{X_1 = 4 \mid X_0 = 1\} \\
&= p_{41} p_{14} = \frac{1}{4}
\end{aligned}
$$

类似可得 $f_{11}^{(3)} = \frac{1}{4}, f_{11}^{(4)} = \frac{1}{4}, f_{11}^{(n)} = 0 (n \geqslant 5)$.

所以 $f_{11} = \sum_{n=1}^{+\infty} f_{11}^{(n)} = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$,故状态1是常返态.

又 $\mu_1 = \sum_{n=1}^{+\infty} n f_{11}^{(n)} = \frac{5}{2} < +\infty$,所以状态1是正常返态.因为各状态相通,依判别准则,所有状态都是正常返态.因而都是遍历的.

4.3.3 Markov 链的状态空间分解

定理 4.8 状态空间中常返状态的全体是一个闭集.

证明 只需证明,若 i 是常返状态,且 $i \to j$,则 $f_{ji} = 1$,从而 $i \leftrightarrow j$,j 是常返状态.事实上,若 $f_{ji} < 1$,则以正概率 $1 - f_{ji} > 0$ 使得从 j 出发不能在有限步内回到 i. 这意味着系统中存在一个正概率,使得从 i 出发不能在有限步内回到 i,从而 $f_{ii} < 1$.

定理 4.9 任意 Markov 链的状态空间 I 可唯一地分解为有限个互不相交的子集 $D, C_1, C_2, \cdots$ 之和,使得

(1) 每个 C_n 是常返状态组成的不可约闭集;

(2) C_n 中的状态同类,具有相同的周期,同为正常返,或同为零常返;

(3) D 中的状态都是非常返状态,从 C_n 中的状态出发不能到 D 中的状态.

证明 设 C 为全体常返状态组成的集合,则 $D = I - C$ 为非常返状态全体,将 C 按互通关系进行分解,得

$$I = D \cup C_1 \cup C_2 \cup \cdots$$

其中每一个 C_n 都是由常返状态组成的不可约闭集. 由于常返性和周期性是类的性质,因此 C_n 中的状态具有相同类型,且 C_n 中的状态不能到达 D 中的状态.

定理 4.10 周期为 d 的不可约 Markov 链,其状态空间 I 可唯一地分解为 d 个互不相交的子集之和,即

$$I = \bigcup_{i=0}^{d-1} I_i, I_i \cap I_j = \varnothing, i \neq j$$

且使得自 I_i 中的任意状态出发,经一步转移必进入 I_{i+1} 中(其中 $I_d = I_0$). (证明略)

以上的状态空间分解定理说明,状态空间的状态可按常返态与非常返态分为两类. 非常返态组成集合 D;常返态组成一个闭集 C,闭集 C 又可按互通关系分为若干个互不相交的基本常返闭集 $C_1, C_2, \cdots$. 一个马尔可夫链,如果从 D 中某个非常返态出发,它或者一直停留在 D 中,或者从某一个时刻进入某个基本常返闭集 C_k,一旦进入,就永不离开;一个马尔可夫链,若从某个常返态出发,必属于某一个常返闭集 C_k,则马尔可夫链将永远停留在这个闭集 C_k 中.

一般地,设 $S \subset I$,如果对 $i \in I$,必存在整数 $n > 0$ 及状态 $j \notin S$,使得 $p_{ij}^{(n)} > 0$,则称 S 为开集. 设 $S \subset I$,则称包含 S 的最小闭集为 S 的闭包,记为 $\bar{S}$,任一状态 i 的闭包 $\overline{\{i\}} = \{i\} \cup \{k : i \to k\}$.

基于以上论述,将状态空间 I 中的状态依 $D, C_1, C_2, \cdots$ 的次序重新排列,则转移概率矩阵具有以下的形式:

$$P = \begin{pmatrix} P_D & P_{D_1} & P_{D_2} & \cdots \\ & P_1 & & \\ & & P_2 & \\ & & & \ddots \end{pmatrix} \begin{matrix} \\ C_1 \\ C_2 \\ \vdots \end{matrix}$$

其中 $P_1, P_2, \cdots$ 均为随机矩阵，它们对应的链是不可约的. 称以上形式的转移概率矩阵为标准形式.

当一个马尔可夫链的状态空间是一个有限集合时，称为**有限马尔可夫链**. 它具有以下性质：

(1) 所有非常返态组成的集合都不是闭集.

(反证法) 若是闭集，则对任何 $i \in D$，有 $\sum_{j\in D} p_{ij}^{(n)} = 1$，但是由非常返性知，$\lim_{n\to+\infty} p_{ij}^{(n)} = 0$，这与 $0 = \sum_{j\in D} \lim_{n\to+\infty} p_{ij}^{(n)} = 1$ 相矛盾.

(2) 有限马尔可夫链没有零常返态.（反证法，读者自己证明）

(3) 有限马尔可夫链必有正常返态.

由转移概率的性质知，总有 $\sum_{j\in I} p_{ij}^{(n)} = 1$. 由于状态空间 I 有限，则求和与取极限运算可交换，于是

$$\lim_{n\to+\infty} \sum_{j\in I} p_{ij}^{(n)} = \sum_{j\in I} \lim_{n\to+\infty} p_{ij}^{(n)} = 1$$

即不可能对一切 j 都有 $\lim_{n\to+\infty} p_{ij}^{(n)} = 0$. 若存在状态 k 使得 $\lim_{n\to+\infty} p_{ik}^{(n)} \neq 0$，则状态 k 即为正常返态.

(4) 有限马尔可夫链的状态空间可分解为

$$I = D + C_1 + \cdots + C_k$$

其中 D 为非常返态集合，$C_i, i = 1, 2, \cdots, k$ 是正常返态集合.

(5) 不可约有限马尔可夫链只有正常返态.

由于有限马尔可夫链是不可约的，因而整个状态空间只能构成唯一一个闭集. 而非常返态集合 D 不是闭集，所以只能是由正常返态组成一个闭集.

【例 4.16】 设马尔可夫链 $\{X_n, n \geq 0\}$ 的状态空间 $I = \{1, 2, \cdots, 6\}$，其转移概率矩阵为

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1/3 & 1/3 & 0 & 1/3 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}$$

试分解此马尔可夫链并求出各状态的周期.

解 状态传递图如图4.7所示.

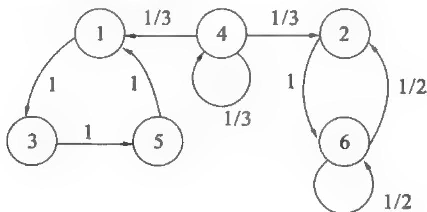


图 4.7

对状态1,有

$$f_{11}^{(1)}=0, f_{11}^{(2)}=0, f_{11}^{(3)}=1, f_{11}^{(n)}=0(n\geqslant 4)$$

故 $f_{11}=1$,状态1为常返态. $\mu_1=\sum_{n=1}^{+\infty}nf_{11}^{(n)}=3$,由状态1生成的基本常返闭集为 $C_1=\{\bar{1}\}=\{1\}\cup\{k:1\to k\}=\{1,3,5\}$.

类似地,状态6是正常返态,$\mu_6=3/2$,由状态6生成的基本闭集 $C_2=\{\bar{6}\}=\{2,6\}$.

$D=\{4\}$ 是非常返集,从而状态空间

$$I=\{4\}\cup\{1,3,5\}\cup\{2,6\}$$

因 $f_{11}^{(3)}>0$,故状态1周期为3,从而 C_1 中状态周期均为3. 又 $f_{66}^{(1)}>0$,故状态6是非周期的,即 C_2 中状态是遍历的. 同样,因 $f_{44}^{(1)}>0$,故状态4也是非周期的.

【例 4.17】 设马尔可夫链的状态空间 $I=\{0,1,2,3\}$,其转移概率矩阵为

$$\boldsymbol{P}=\begin{pmatrix}1/2 & 1/2 & 0 & 0\\ 1/2 & 1/2 & 0 & 0\\ 1/4 & 1/4 & 1/4 & 1/4\\ 0 & 0 & 0 & 1\end{pmatrix}$$

求状态空间的分解.

解 状态传递图如图4.8所示.

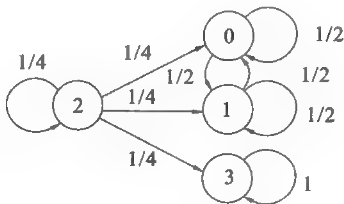


图 4.8

由于状态3不可能到达任何其他状态,所以是常返态. 状态2可到达0,1,3三个状态,但从0,1,3三个状态都不能到达状态2,但0与1两状态相通,构成一个常

返态闭集. 又 $\sum_{n=1}^{+\infty} f_{22}^{(n)} = \frac{1}{4} < 1$,所以状态 2 为非常返态. 于是状态空间分解为 $I = \{2\} + \{0,1\} + \{3\}$.

4.3.4 类之间的转移和赌徒输光问题

首先考虑如下问题:设 j 是一给定的常返状态,D 是由全部非常返状态组成的集合. 对 $i \in D$,如何计算 f_{ij},即过程从 i 出发迟早到达 j 的概率.

定理 4.11 若 j 是常返的,则概率组 $\{f_{ij}, i \in D\}$ 满足

$$f_{ij} = \sum_{k \in D} p_{ik} f_{kj} + \sum_{k \in C} p_{ik}$$

其中 C 是所有与 j 相通的状态的集合.

证明

$$\begin{aligned} f_{ij} &= P\{\exists n < +\infty, X_n = j \mid X_0 = i\} \\ &= \sum_{k \in S} P\{\exists n < +\infty, X_n = j \mid X_0 = i, X_1 = k\} P\{X_1 = k \mid X_0 = i\} \\ &= \sum_{k \in D} p_{ik} f_{kj} + \sum_{k \in C} p_{ik} f_{kj} + \sum_{\substack{k \notin C \\ k \notin D}} p_{ik} f_{kj} \\ &= \sum_{k \in D} p_{ik} f_{kj} + \sum_{k \in C} p_{ik} \end{aligned}$$

其中第二个等式是利用 C-K 方程,最后一个等式是利用若 k 和 j 互通,则 $f_{kj} = 1$.

【例 4.18】 (赌徒输光问题)考虑一赌徒,在每局赌博中他以概率 p 赢 1 元,以概率 $q = 1 - p$ 输 1 元,如果他输光或获得 N 元,则赌局结束. 假定各局赌博是独立的,赌徒开始有 i 元,问他的赌金到达 0(输光)之前达到 N 元的概率是多少?

解 以 X_n 记赌徒在时刻 n 的赌金,则 $\{X_n, n = 0, 1, \cdots\}$ 是 Markov 链,其转移概率为

$$p_{00} = p_{NN} = 1$$

$$p_{i,i+1} = p, \quad p_{i,i-1} = 1 - p, \quad i = 1, 2, \cdots, N-1$$

此 Markov 链可分为三类:$\{0\}$、$\{N\}$、$\{1, 2, \cdots, N-1\}$,其中第一和第二类是常返的,第三类是非常返的.

记 $f_i = f_{iN}$,表示赌徒从 i 元的赌本开始,最终达到 N 元的概率. 根据命题,有

$$f_i = p f_{i+1} + q f_{i-1}, i = 1, 2, \cdots, N-1$$

由于 $p + q = 1$,故

$$f_{i+1} - f_i = \frac{q}{p}(f_i - f_{i-1})$$

由于 $f_0 = 0$,从上式可见

$$f_2 - f_1 = \frac{q}{p}(f_1 - f_0) = \left(\frac{q}{p}\right) f_1$$

$$f_3 - f_2 = \frac{q}{p}(f_2 - f_1) = \left(\frac{q}{p}\right)^2 f_1$$

$$\vdots$$

$$f_i - f_{i-1} = \frac{q}{p}(f_{i-1} - f_{i-2}) = \left(\frac{q}{p}\right)^{i-1} f_1$$

$$\vdots$$

$$f_N - f_{N-1} = \frac{q}{p}(f_{N-1} - f_{N-2}) = \left(\frac{q}{p}\right)^{N-1} f_1$$

将前 $i-1$ 个方程相加得

$$f_i - f_1 = f_1[(q/p) + (q/p)^2 + \cdots + (q/p)^{i-1}]$$

即

$$f_i = \begin{cases} \dfrac{1-(q/p)^i}{1-(q/p)} f_1, & q/p \neq 1 \\ i f_1, & q/p = 1 \end{cases}$$

利用 $f_N = 1$ 得

$$f_1 = \begin{cases} \dfrac{1-(q/p)}{1-(q/p)^N}, & p \neq 1/2 \\ 1/N, & p = 1/2 \end{cases}$$

因此，

$$f_i = \begin{cases} \dfrac{1-(q/p)^i}{1-(q/p)^N}, & p \neq 1/2 \\ i/N, & p = 1/2 \end{cases}$$

注意到，$N \to +\infty$ 时，

$$f_i = \begin{cases} 1-(q/p)^i, & p \neq 1/2 \\ 0, & p \leqslant 1/2 \end{cases}$$

因此由概率的连续性知，在与有无穷赌本的对手的赌博中，当 $p > 1/2$ 时，赌徒的赌本以一正概率趋于无穷，而当 $p \leqslant 1/2$ 时，将以概率 1 输光.

4.4 Markov 链的极限定理与平稳分布

本节主要对 $p_{ij}^{(n)}$ 的极限考虑如下几个问题：一是当 $n \to +\infty$ 时，$p_{ij}^{(n)}$ 的极限是否存在？二是若存在，$\pi_{ij} = \lim\limits_{n\to+\infty} p_{ij}^{(n)}$ 是否与初始状态 i 有关？若不存在，$\lim\limits_{n\to+\infty} \dfrac{1}{n}\sum\limits_{k=1}^{n} p_{ij}^{(k)}$ 是否存在？是否与初始状态 i 有关？这就与马尔可夫链的所谓平稳分布有密切关系. 三是对于 Markov 链，中心极限定理和强大数定律是否成立？

4.4.1 $p_{ij}^{(n)}$ 的极限性质

首先来看两个例子.

【例 4.19】 设 Markov 链的转移概率矩阵为

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \quad 0 < p,q < 1$$

试计算 $\lim\limits_{n\to+\infty} p_{ij}^{(n)}, \forall i,j \in I$.

解 由于 $p_{ij}^{(n)}$ 是矩阵 $P^{(n)} = P^n$ 的元素,因此,只需计算 $\lim\limits_{n\to\infty} P^n$. 根据线性代数的知识,矩阵 P 可以分解为

$$P = QDQ^{-1}$$

其中

$$Q = \begin{pmatrix} 1 & -p \\ 1 & q \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 1-p-q \end{pmatrix}$$

$$Q^{-1} = \begin{bmatrix} \dfrac{q}{p+q} & \dfrac{p}{p+q} \\ -\dfrac{1}{p+q} & \dfrac{1}{p+q} \end{bmatrix}$$

从而

$$P^n = (QDQ^{-1})^n = Q \begin{pmatrix} 1 & 0 \\ 0 & 1-p-q \end{pmatrix}^n Q^{-1}$$

$$= \begin{bmatrix} \dfrac{q+p(1-p-q)^n}{p+q} & \dfrac{p-p(1-p-q)^n}{p+q} \\ \dfrac{q-p(1-p-q)^n}{p+q} & \dfrac{p+p(1-p-q)^n}{p+q} \end{bmatrix}$$

由于 $|1-p-q| < 1$,所以上式的极限为

$$\lim_{n\to+\infty} P^n = \begin{bmatrix} \dfrac{q}{p+q} & \dfrac{p}{p+q} \\ \dfrac{q}{p+q} & \dfrac{p}{p+q} \end{bmatrix}$$

可以看出 $\pi_{ij} = \lim\limits_{n\to+\infty} p_{ij}^{(n)}$ 存在,且与初始状态 i 无关,只与状态 j 有关.

【例 4.20】 考虑简单随机游动,转移概率可以写为

$$p_{i,i+1} = 1 - p_{i,i-1} = p$$

容易证明

$$p_{00}^{(2n)} = \binom{2n}{n} p^n (1-p)^n = \frac{(2n)!}{n!n!} [p(1-p)]^n \sim \frac{[4p(1-p)]^n}{\sqrt{n\pi}}$$

当 $p=\frac{1}{3}$ 时，0 是非常返状态，且有

$$\lim_{n\to+\infty} p_{00}^{(2n)} = \lim_{n\to+\infty} \frac{\left(4\times\frac{1}{3}\times\frac{2}{3}\right)^n}{\sqrt{n\pi}} = 0$$

当 $p=\frac{1}{2}$ 时，0 是零常返状态，且有

$$\lim_{n\to+\infty} p_{00}^{(2n)} = \lim_{n\to+\infty} \frac{\left(4\times\frac{1}{2}\times\frac{1}{2}\right)^n}{\sqrt{n\pi}} = 0$$

由此可以看出，无论状态 0 是非常返状态还是零常返状态，都有 $\lim\limits_{n\to+\infty} p_{00}^{(n)}=0$. 事实上，这个结论是具有普遍意义的.

定理 4.12 若状态 j 是非常返或零常返状态，则 $\pi_{ij}=\lim\limits_{n\to+\infty} p_{ij}^{(n)}=0, \forall i\in I$.

证明 由公式

$$p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$$

对 $N<n$ 有

$$\sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)} \leqslant \sum_{l=1}^{N} f_{ij}^{(l)} p_{jj}^{(n-l)} + \sum_{l=N+1}^{n} f_{ij}^{(l)}$$

先固定 N，令 $n\to+\infty$，由于 $p_{jj}^{(n)}\to 0$，所以上式右端第一项趋于 0. 再令 $N\to+\infty$，上式右端的第二项因 $\sum\limits_{l=1}^{+\infty} f_{ij}^{(n)}\leqslant 1$ 而趋于 0，故

$$\lim_{n\to+\infty} p_{ij}^{(n)} = 0$$

定理 4.13 若状态 j 是周期为 d 的常返状态，则

$$\lim_{n\to+\infty} p_{jj}^{(nd)} = \frac{d}{\mu_j}$$

当 $\mu_j\to+\infty$ 时，$\frac{d}{\mu_j}=0$.

证明 设 $X_0=j$，以 $N_j(t)$ 记到时刻 t 为止转移到 j 的次数. 由于 j 是常返态，则一旦转移到 j，在概率上就重新开始. 故 $\{N_j(t),t\geqslant 0\}$ 是一个到达时间间隔分布为 $\{f_{jj}^{(n)},n\geqslant 1\}$ 的更新过程. 从而

$$p_{jj}^{(nd)} = P\{X_{nd}=j \mid X_0=j\} = P\{\text{在 } nd \text{ 处更新}\}$$

由 Blackwell 更新定理有，

$$\lim_{n\to+\infty} p_{jj}^{(nd)} = \frac{d}{\mu_j}$$

推论 4.3 设 i 是常返状态，则

$$i\text{ 是零常返状态} \Leftrightarrow \lim_{n\to+\infty} p_{ii}^{(n)} = 0$$

证明 若 i 为零常返状态，则 $\mu_i=+\infty$，从而 $\lim\limits_{n\to+\infty} p_{ii}^{(nd)}=0$. 而当 n 不是 d 的整数倍时，$p_{ii}^{(n)}=0$. 故 $\lim\limits_{n\to+\infty} p_{ii}^{(n)}=0$.

反过来，若 $\lim\limits_{n\to+\infty} p_{ii}^{(n)}=0$，且 i 是正常返状态，则 $\mu_i<+\infty$，根据定理可知，$\lim\limits_{n\to+\infty} p_{ii}^{(nd)}=\dfrac{d}{\mu_i}>0$，这与 $\lim\limits_{n\to\infty} p_{ii}^{(n)}=0$ 矛盾.

推论 4.4 零(正)常返是类的性质.

证明 设 $i\leftrightarrow j$ 为常返状态，且 i 为零常返状态，则

$$p_{ii}^{(n)}\to 0$$

且存在 n,m 使得 $p_{ij}^{(n)}>0, p_{ji}^{(m)}>0$. 由 C-K 方程得

$$p_{ii}^{(n+m+l)}\geqslant p_{ij}^{(n)}p_{jj}^{(l)}p_{ji}^{(m)}\geqslant 0$$

令 $l\to+\infty$，对上式取极限得

$$0\geqslant \lim_{l\to+\infty} p_{jj}^{(l)}\geqslant 0$$

故 j 也为零常返状态.

反之，由 j 为零常返状态也可推出 i 为零常返状态，即证明了 i,j 同为零常返或正常返态.

推论 4.5 有限状态的 Markov 链，不可能全为非常返状态，也不可能有零常返状态，从而不可约的有限状态的 Markov 链是正常返的.

证明 设状态空间 $I=\{1,2,\cdots,N\}$，若全部状态为非常返状态，对状态 $i\to j$，有 $p_{ij}^{(n)}\to 0$. 若 i 不可达 j，则对任意 n，$p_{ij}^{(n)}=0$. 因此，当 $n\to+\infty$ 时，有 $\lim\limits_{n\to+\infty}\sum\limits_{j=1}^{N} p_{ij}^{(n)}=0$，这与 $\sum\limits_{j=1}^{N} p_{ij}^{(n)}=1$ 矛盾.

推论 4.6 若 Markov 链有一个零常返状态，则必有无限个零常返状态.

证明 若 I 中有零常返状态，设为 i，令 C 为所有与 i 互通的状态的集合，则由 Markov 链的状态分解定理可知，C 是一个封闭子集，对任意 n，$\sum\limits_{j\in C} p_{ij}^{(n)}=1$. 对于 C 中的任意状态 j，也是零常返状态，所以

$$\lim_{n\to+\infty} p_{ij}^{(n)}=0$$

从而

$$\lim_{n\to+\infty}\sum_{j\in C} p_{ij}^{(n)}=0$$

与 $\sum\limits_{j\in C} p_{ij}^{(n)}=1$ 矛盾.

定理 4.13 的结论给出了非常返状态或零常返状态的渐近分布，但是当 j 是正常返状态时，$\lim\limits_{n\to+\infty} p_{ij}^{(n)}$ 不一定存在，即使存在也可能与 i 有关. 因此，下面借助于研究

$p_{ij}^{(nd)}(d \geqslant 1)$ 和 $\frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)}$ 的极限性质，来讨论 j 是正常返状态时 $\lim_{n\to+\infty} p_{ij}^{(n)}$ 的性态.

令 $f_{ij}(r) = \sum_{m=0}^{+\infty} f_{ij}^{(md+r)}$，$0 \leqslant r \leqslant d-1$ 表示从状态 i 出发，在时刻 $n = r\text{mod}(d)$ 首次达到状态 j 的概率，显然

$$\sum_{r=0}^{d-1} f_{ij}(r) = \sum_{m=0}^{+\infty}\sum_{r=0}^{d-1} f_{ij}^{(md+r)} = \sum_{m=0}^{+\infty} f_{ij}^{(m)} = f_{ij}$$

定理 4.14 若 j 是正常返状态，且周期为 d，则对任意 i 及 $0 \leqslant r \leqslant d-1$，有

$$\lim_{n\to+\infty} p_{ij}^{(nd+r)} = f_{ij}(r)\frac{d}{\mu_j}$$

证明 因为当 $n \neq kd$ 时，$p_{jj}^{(n)} = 0$，所以

$$p_{ij}^{(nd+r)} = \sum_{k=0}^{nd+r} f_{ij}^{(k)} p_{jj}^{(nd+r-k)} = \sum_{m=0}^{n} f_{ij}^{(md+r)} p_{jj}^{(n-m)d}$$

于是，对 $1 \leqslant N < n$ 有

$$\sum_{m=0}^{N} f_{ij}^{(md+r)} p_{jj}^{(n-m)d} \leqslant p_{ij}^{(nd+r)} \leqslant \sum_{m=0}^{N} f_{ij}^{(md+r)} p_{jj}^{(n-m)d} + \sum_{m=N+1}^{n} f_{ij}^{(md+r)}$$

在上式中，先令 $n \to +\infty$，再令 $N \to +\infty$，由定理 4.13 得

$$f_{ij}(r)\frac{d}{\mu_j} \leqslant \lim_{n\to+\infty} p_{ij}^{(nd+r)} \leqslant f_{ij}(r)\frac{d}{\mu_j}$$

故

$$\lim_{n\to+\infty} p_{ij}^{(nd+r)} = f_{ij}(r)\frac{d}{\mu_j}$$

推论 4.7 设不可约的、正常返的、周期为 d 的 Markov 链，其状态空间为 I，则对任何状态 $i \to j$，$j \in I$ 有

$$\lim_{n\to+\infty} p_{ij}^{(nd)} = \begin{cases} \dfrac{d}{\mu_j}, & \text{若 } i \text{ 与 } j \text{ 同属于子集 } I_k \\ 0, & \text{其他} \end{cases}$$

其中 $I = \bigcup_{k=1}^{d-1} I_k$. 特别地，当 $d = 1$ 时，则 $\forall i,j \in I$ 有

$$\lim_{n\to+\infty} p_{ij}^{(n)} = \frac{1}{\mu_j}$$

这是不可约的、正常返的、周期 $d = 1$ 的 Markov 链的直观含义.

证明 在定理 4.14 中取 $r = 0$ 得

$$\lim_{n\to+\infty} p_{ij}^{(nd)} = f_{ij}(0)\frac{d}{\mu_j}$$

其中 $f_{ij}(0) = \sum_{m=0}^{+\infty} f_{ij}^{(md)}$.

若 i 与 j 不在同一个 I_k 中，则由定理知，$p_{ij}^{(nd)} = 0$，从而 $f_{ij}^{(nd)} = 0$，于是 $f_{ij} = 0$.

若 i 与 j 在同一个 I_k 中，则当 n 不是 d 的整数倍时，$p_{ij}^{(n)}=0$，从而 $f_{ij}^{(n)}=0$，因此有

$$f_{ij}(0)=\sum_{m=0}^{+\infty}f_{ij}^{(md)}=\sum_{m=0}^{+\infty}f_{ij}^{(m)}=f_{ij}=1$$

虽然一般情况下，$\lim\limits_{n\to+\infty}p_{ij}^{(n)}$ 不一定存在，但是可以得到平均值 $\lim\limits_{n\to+\infty}\frac{1}{n}\sum\limits_{k=1}^{n}p_{ij}^{(k)}$ 的极限定理.

定理 4.15　对于任意状态 $i,j\in I$ 有

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n}p_{ij}^{(k)}=\begin{cases}0, & j\text{ 为非常返状态或零常返状态}\\ \dfrac{f_{ij}}{\mu_j}, & j\text{ 为正常返状态}\end{cases}$$

证明　借助下面的结论：假设有非负数列 $\{a_n\}$ 的 d 个子列 $\{a_{kd+s}\}$，$s=0,1,\cdots,d-1$，如果对每一个 s 存在极限 $\lim\limits_{k\to+\infty}a_{kd+s}=b_s$，则有

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n}a_k=\frac{1}{d}\sum_{s=0}^{d-1}b_s$$

若 j 是非常返或零常返状态，由定理 4.12 知 $\lim\limits_{n\to+\infty}p_{ij}^{(n)}=0$，所以

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n}p_{ij}^{(k)}=0$$

若 j 为正常返状态且周期为 d，则令 $a_{kd+s}=p_{ij}^{(kd+s)}$，然后利用定理 4.14 得 $b_s=f_{ij}(s)\dfrac{d}{\mu_j}$. 由上述结论可知

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n}p_{ij}^{(k)}=\frac{1}{d}\sum_{s=0}^{d-1}f_{ij}(s)\frac{d}{\mu_j}=\frac{1}{\mu_j}\sum_{s=0}^{d-1}f_{ij}(s)=\frac{f_{ij}}{\mu_j}$$

推论 4.8　如果 $\{X_n\}$ 是不可约的、常返状态的 Markov 链，则对任意状态 $i,j\in I$ 有

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n}p_{ij}^{(k)}=\frac{1}{\mu_j}$$

当 $\mu_j=+\infty$ 时，可以认定 $\dfrac{1}{\mu_j}=0$.

定理 4.15 及推论 4.8 指出，j 为正常返状态时，$\lim\limits_{n\to+\infty}p_{ij}^{(n)}$ 不一定存在，但是 $\lim\limits_{n\to+\infty}\frac{1}{n}\sum\limits_{k=1}^{n}p_{ij}^{(k)}$ 是存在的，当马氏链不可约时，其极限与 i 无关.

注：本小节研究当 $n\to+\infty$ 时转移概率 $p_{ij}^{(n)}$ 的极限性质，即研究 $P\{X_n=j\mid X_0=i\}$ 的极限分布，主要包含两个问题，一是 $\lim\limits_{n\to+\infty}p_{ij}^{(n)}$ 是否存在？二是若存在，$\lim\limits_{n\to+\infty}p_{ij}^{(n)}$ 是否与初始状态 i 有关？在马氏链理论中，这一类问题的定理统称为遍历性

定理. 研究转移概率 $p_{ij}^{(n)}$ 的遍历性给解决实际问题带来了许多方便.

若对于 $i,j \in I$,存在不依赖于 i 的极限 $\lim\limits_{n\to+\infty} p_{ij}^{(n)} = p_j > 0$,则称齐次马氏链具有遍历性. 一个不可约的齐次马氏链,若它的状态是非周期正常返状态,则这个状态就具有遍历性,从而它是一个遍历链. 具有遍历性的马氏链,无论系统从哪个状态出发,当转移的步数 n 充分大时,转移到状态 j 的概率都近似于 p_j. 在实际运用中,当 n 充分大时,可以用 p_j 作为 $p_{ij}^{(n)}$ 的近似值.

4.4.2 平稳分布与极限分布

我们知道,判别一个不可约的、非周期的、常返态的马尔可夫链是否为遍历的,可以通过讨论 $\lim\limits_{n\to+\infty} p_{ij}^{(n)}$ 来解决. 但是,在实际问题中求 $\lim\limits_{n\to+\infty} p_{ij}^{(n)}$ 是比较困难的,所以,通常借助于研究马氏链的平稳分布是否存在,来判别齐次马尔可夫链是否为遍历链.

定义 4.13 一个定义在状态空间上的概率分布 $\boldsymbol{\pi}=\{\pi_1,\pi_2,\cdots,\pi_i,\cdots\}$ 称为**马氏链的平稳分布**,如有 $\boldsymbol{\pi}=\boldsymbol{\pi}\boldsymbol{P}$,即 $\forall j \in I$,有 $\pi_j = \sum\limits_{i\in I}\pi_i p_{ij}$.

平稳分布也称为马氏链的不变概率测度. 对于一个平稳分布 $\boldsymbol{\pi}$,显然有

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}\boldsymbol{P}^2 = \cdots = \boldsymbol{\pi}\boldsymbol{P}^n$$

注:理解"平稳"的含义可借助于下列命题:如果 $\{\pi_k\}$ 是马尔可夫链 $\{X_n, n\geqslant 0\}$ 的平稳分布,对每一个 k 都有 $P\{X_0 = k\} = \pi_k$,则对一切 $n\geqslant 1$,有 $P\{X_n = k\} = \pi_k$,且对任意正整数 n 及状态 $j_v, v=0,1,\cdots,l$ 有

$$P\{X_{n+v} = j_v, 0\leqslant v\leqslant l\} = P\{X_v = j_v, 0\leqslant v\leqslant l\}$$

定理 4.16 设 $\{X_n, n\geqslant 0\}$ 是一马氏链,则 $\{X_n, n\geqslant 0\}$ 为平稳过程的充分必要条件是 $\boldsymbol{\pi}(0) = \{\pi_i(0), i\in I\}$ 是平稳分布,即有 $\boldsymbol{\pi}(0) = \boldsymbol{\pi}(0)\boldsymbol{P}$.

证明 充分性:记 $\boldsymbol{\pi}(0) \triangleq \boldsymbol{\pi}$,则有

$$\boldsymbol{\pi}(1) = \boldsymbol{\pi}(0)\boldsymbol{P} = \boldsymbol{\pi}$$

$$\boldsymbol{\pi}(2) = \boldsymbol{\pi}(1)P = \boldsymbol{\pi}(0)P^2 = \boldsymbol{\pi}$$

$$\vdots$$

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(n-1)\boldsymbol{P} = \cdots = \boldsymbol{\pi}$$

因此,对于 $\forall i_k \in I, t_k \in \mathbf{N}, n\geqslant 1, 1\leqslant k\leqslant n, t\in \mathbf{N}$,有

$$\begin{aligned}
P\{X_{t_1} = i_1, X_{t_2} = i_2, \cdots, X_{t_n} = i_n\} &= \pi_{i_1}(t_1) p_{i_1 i_2}^{(t_2-t_1)} \cdots p_{i_{n-1} i_n}^{(t_n-t_{n-1})} \\
&= \pi(t_1+t) p_{i_1 i_2}^{(t_2-t_1)} \cdots p_{i_{n-1} i_n}^{(t_n-t_{n-1})} \\
&= P\{X_{t_1+t} = i_1, X_{t_2+t} = i_2, \cdots, X_{t_n+t} = i_n\}
\end{aligned}$$

所以 $\{X_n, n\geqslant 0\}$ 是严平稳过程.

必要性：由于$\{X_n, n \geqslant 0\}$是平稳过程，因此有

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(n-1) = \cdots = \boldsymbol{\pi}(0)$$

又由$\boldsymbol{\pi}(1) = \boldsymbol{\pi}(0)P$得

$$\boldsymbol{\pi}(0) = \boldsymbol{\pi}(0)P$$

即$\boldsymbol{\pi}(0)$是平稳分布.

由定义4.13和定理4.14不难得到如下结果：

定理4.17　不可约的遍历马尔可夫链恒有唯一的平稳分布$\left\{\pi_i = \dfrac{1}{\mu_i}, i \in I\right\}$，且$\pi_j = \lim\limits_{n\to+\infty} p_{ij}^{(n)}$.

定义4.14　若$\lim\limits_{n\to+\infty} \pi_j(n) = \lim\limits_{n\to+\infty} P\{X_n = j\} = \pi_j^*, j \in I$存在，则称$\boldsymbol{\pi}^* = \{\pi_1^*, \cdots, \pi_j^*, \cdots\}$为马氏链的极限分布.

定理4.18　不可约的、非周期的马尔可夫链是正常返状态的充分必要条件是它存在平稳分布，且此平稳分布就是极限分布$\{\dfrac{1}{\mu_j}, j \in I\}$.

证明　充分性：设存在平稳分布$\boldsymbol{\pi} = \{\pi_1, \pi_2, \cdots, \pi_j, \cdots\}$，由此有

$$\boldsymbol{\pi} = \boldsymbol{\pi P} = \boldsymbol{\pi P}^2 = \cdots = \boldsymbol{\pi P}^n$$

即

$$\pi_j = \sum_{i\in I} \pi_i p_{ij}^{(n)}$$

由于$\pi_j \geqslant 0, j \in I, \sum\limits_{j\in I} \pi_j = 1$.

由控制收敛定理，有

$$\pi_j = \lim_{n\to+\infty} \sum_{i\in I} \pi_i p_{ij}^{(n)} = \sum_{i\in I} \pi_i \lim_{n\to+\infty} p_{ij}^{(n)} = \left(\sum_{i\in I} \pi_i\right)\frac{1}{\mu_j} = \frac{1}{\mu_j}$$

因为

$$\sum_{j\in I} \pi_j = \sum_{j\in I} \frac{1}{\mu_j} = 1$$

于是至少存在一个$\pi_l = \dfrac{1}{\mu_l} > 0$，从而

$$\lim_{n\to+\infty} p_{il}^{(n)} = \frac{1}{\mu_l} > 0$$

即有$\mu_l < +\infty$，故l为正常返状态. 由不可约性可知，整个马尔可夫链是正常返的，且有

$$\pi_j = \frac{1}{\mu_j} > 0, j \in I$$

必要性：由于马尔可夫链是正常返非周期链，即为遍历马尔可夫链，由以上的

定理立即可得结果,且有

$$\pi_j = \pi_j^* = \frac{1}{\mu_j}, j \in I$$

由此定理可知,对于不可约的遍历马尔可夫链,其极限分布 $\boldsymbol{\pi}^* = \boldsymbol{\pi}$ 存在,且就是等于平稳分布.

【例 4.21】 设马尔可夫链的状态空间为 $I = \{0,1,2\}$,其一步转移概率矩阵为

$$\boldsymbol{P} = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}$$

求其相应的极限分布.

解 设其极限分布 $\boldsymbol{\pi} = (\pi_0, \pi_1, \pi_2)$,由 $\boldsymbol{\pi} = \boldsymbol{\pi P}$ 得方程组

$$\begin{cases} 0.5\pi_0 + 0.3\pi_1 + 0.2\pi_2 = \pi_0 \\ 0.4\pi_0 + 0.4\pi_1 + 0.3\pi_2 = \pi_1 \\ 0.1\pi_0 + 0.3\pi_1 + 0.5\pi_2 = \pi_2 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases}$$

解方程组,得 $\pi_0 = 21/62, \pi_1 = 23/62, \pi_2 = 18/62$. 即不论其初始分布如何,在经过一段时间以后,有 21/62 的时间过程处于状态 0,有 23/62 的时间过程处于状态 1,有 18/62 的时间过程处于状态 2.

【例 4.22】 设马尔可夫链的状态空间为 $I = \{1,2,3,4\}$,其转移概率矩阵为

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/2 \end{bmatrix}$$

讨论 $\lim\limits_{n\to+\infty} p_{i1}^{(n)}$.

解 状态 1 和状态 2 都是吸收态,都是正常返非周期的基本常返闭集,而 $N = \{3,4\}$ 是非常返集,有

$$p_{11}^{(n)} = 1, p_{21}^{(n)} = 0, p_{31}^{(n)} = 1/3$$

$$p_{41}^{(n)} = \sum_{l=1}^{n} f_{41}^{(l)} \cdot p_{11}^{(n-l)} = \sum_{l=1}^{n} f_{41}^{(l)}$$

$$= \frac{1}{4} + \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{4} + \cdots + \frac{1}{2^{n-1}}\cdot\frac{1}{4} = \frac{1}{2} - \frac{1}{2^{n+1}}$$

说明 $\lim\limits_{n\to+\infty} p_{i1}^{(n)}$ 存在,但与 i 有关.

【例 4.23】 设一马尔可夫链的转移概率矩阵为

$$P=\begin{pmatrix}0 & 1/2 & 0 & 1/2\\ 1/2 & 0 & 1/2 & 0\\ 0 & 1/2 & 0 & 1/2\\ 1/2 & 0 & 1/2 & 0\end{pmatrix}$$

讨论此马尔可夫链的遍历性.

解　因为

$$P^{(2)}=\begin{pmatrix}1/2 & 0 & 1/2 & 0\\ 0 & 1/2 & 0 & 1/2\\ 1/2 & 0 & 1/2 & 0\\ 0 & 1/2 & 0 & 1/2\end{pmatrix},\cdots$$

当 n 为偶数时,$P^{(n)}=P^{(2)}$. 所以,对任一固定的状态 $j(j=1,2,3,4)$,极限 $\lim\limits_{n\to+\infty}p_{ij}^{(n)}$ 都不存在. 故此马尔可夫链不具有遍历性.

习　题　4

1. 独立地重复抛掷一枚硬币,每次抛掷出现正面的概率为 p. 对于 $n\geqslant2$,令 $X_n=0,1,2,3$,这些值分别对应于第 $n-1$ 次和第 n 次抛掷的结果为(正,正),(正,反),(反,正),(反,反). 求马尔可夫链 $\{X_n,n=0,1,2,\cdots\}$ 的一步和两步转移概率矩阵.
2. 设 $\{X(t),t\in T\}$ 为随机过程,且
$$X_1=X(t_1),X_2=X(t_2),\cdots,X_n=X(t_n),\cdots$$
为独立同分布随机变量序列,令
$$Y_0=0,Y_1=Y(t_1)=X_1,Y_n+cY_{n-1}=X_n,n\geqslant2$$
试证:$\{Y_n,n\geqslant0\}$ 是马尔可夫链.
3. 设今日有雨且明日也有雨的概率为 0.7,今日无雨而明日有雨的概率为 0.5. 求星期一有雨,星期三也有雨的概率.
4. 已知随机游动的转移概率矩阵为
$$P=\begin{pmatrix}0.5 & 0.5 & 0\\ 0 & 0.5 & 0.5\\ 0.5 & 0 & 0.5\end{pmatrix}$$
求其三步转移概率矩阵 $P^{(3)}$ 及初始分布为
$$P\{X_0=1\}=P\{X_0=2\}=0,P\{X_0=3\}=1$$
时,经三步转移后处于状态 3 的概率.
5. 已知本月销售状态的初始分布和转移概率矩阵分别如下:

(1)$P^T(0)=(0.4,0.2,0.4),P=\begin{pmatrix}0.8 & 0.1 & 0.1\\ 0.1 & 0.7 & 0.2\\ 0.2 & 0.2 & 0.6\end{pmatrix}$;

(2)$P^T(0)=(0.2,0.2,0.3,0.3)$，$P=\begin{pmatrix}0.7&0.1&0.1&0.1\\0.1&0.6&0.2&0.1\\0.1&0.1&0.6&0.2\\0.1&0.1&0.2&0.6\end{pmatrix}$，

求接下来一个月、两个月的销售状态分布.

6. 讨论下列转移概率矩阵的马尔可夫链的状态分类.

(1)$P=\begin{pmatrix}0.2&0.3&0.5&0&0\\0.7&0.3&0&0&0\\0&1&0&0&0\\0&0&0&0.4&0.6\\0&0&0&1&0\end{pmatrix}$；

(2)$P=\begin{pmatrix}0&0&1&0\\1&0&0&0\\0.3&0.7&0&0\\0.6&0.2&0.2&0\end{pmatrix}$；

(3)$P=\begin{pmatrix}1&0&\cdots&\cdots&\cdots&\cdots&\cdots&0\\q&r&p&0&\cdots&\cdots&\cdots&0\\0&q&r&p&0&\cdots&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&\cdots&\cdots&\cdots&0&q&r&p\\0&\cdots&\cdots&\cdots&\cdots&\cdots&0&1\end{pmatrix}$.

其中 $q+r+p=1$，$I=\{0,1,\cdots,n\}$.

7. 设马尔可夫链的状态空间为 $I=\{1,2,\cdots,7\}$，转移概率矩阵为

$$P=\begin{pmatrix}0.4&0.2&0.1&0&0.1&0.1&0.1\\0.1&0.3&0.2&0.2&0&0.1&0.1\\0&0&0.6&0.4&0&0&0\\0&0&0.4&0&0.6&0&0\\0&0&0.2&0.5&0.3&0&0\\0&0&0&0&0&0.3&0.7\\0&0&0&0&0&0.8&0.2\end{pmatrix}$$

求状态的分布及各常返闭集的平稳分布.

8. 设马尔可夫链的转移概率矩阵为

(1)$\begin{pmatrix}0.5&0.5\\1/3&2/3\end{pmatrix}$；(2)$\begin{pmatrix}p_1&q_1&0\\0&p_2&q_2\\q_3&0&p_3\end{pmatrix}$.

计算 $f_{11}^{(n)}, f_{12}^{(n)}, n=1,2,3$.

9. 设马尔可夫链的转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots \\ q_1 & 0 & p_1 & 0 & \cdots & \cdots \\ 0 & q_2 & 0 & p_2 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

求它的平稳分布.

10. 将2个红球4个白球任意地分别放入甲乙两个盒子中,每个盒子放3个.现从每个盒子中各取一球,交换后放回盒子中(甲盒内取出的球放入乙盒中,乙盒内取出的球放入甲盒中).以 $X(n)$ 记第 n 次交换后甲盒中的红球数,则 $\{X(n), n\geqslant 0\}$ 为一齐次马尔可夫链,试求:

(1) 一步转移概率矩阵;

(2) 证明 $\{X(n), n\geqslant 0\}$ 是遍历链;

(3) 求 $\lim\limits_{n\to+\infty} P_{ij}^{(n)}, j=0,1,2$.

11. 转移矩阵称为双随机的,若对一切 j, $\sum\limits_{i=0}^{+\infty} p_{ij}=1$. 设一个双随机转移矩阵的马氏链有 n 个状态,且是遍历的,求它的极限概率.

12. 设 $\{X(n), n\geqslant 1\}$ 为非周期不可约马尔可夫链,状态空间为 I,若对一切 $j\in I$,其一步转移概率矩阵满足条件: $\sum\limits_{i\in I} p_{ij}=1$,试证:

(1) 对一切 j, $\sum\limits_{i\in I} p_{ij}=1$;

(2) 若状态空间 $I=\{1,2,\cdots,m\}$,计算各状态的平均返回时间.

13. 设河流每天的BOD(生物耗氧量)浓度为齐次马氏链,状态空间 $I=\{1,2,3,4\}$ 是按BOD浓度为极低、低、中、高分别表示的,其一步转移概率矩阵(以1 d天为单位)为

$$\boldsymbol{P}=\begin{pmatrix} 0.5 & 0.4 & 0.1 & 0 \\ 0.2 & 0.5 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.6 & 0.1 \\ 0 & 0.2 & 0.4 & 0.4 \end{pmatrix}$$

若BOD浓度为高,则称河流处于污染状态.

(1) 证明该马氏链是遍历的;

(2) 求该马氏链的平稳分布;

(3) 河流再次达到污染的平均时间 μ_4.

第 5 章　连续时间马尔可夫链

5.1　连续时间马尔可夫链的概念

定义 5.1　设 $X=\{X(t),t\geqslant 0\}$ 是取值于状态空间 I 的随机过程，I 是有限或无限可列的，如果对于任意的正整数 n，任意的 $0\leqslant t_1<t_2<\cdots<t_n<t_{n+1}$，及任意的状态 $i_1,i_2,\cdots,i_n,i_{n+1}\in I$，均有

$$\begin{aligned}P\{X(t_{n+1})=i_{n+1}\mid X(t_1)=i_1,X(t_2)=i_2,\cdots,X(t_n)=i_n\}\\=P\{X(t_{n+1})=i_{n+1}\mid X(t_n)=i_n\}\end{aligned}$$

则称 $\{X(t),t\geqslant 0\}$ 是连续时间的 Markov 链，也称时间连续、状态离散的 Markov 过程或可数状态的 Markov 过程.

对于连续时间的 Markov 链，有

$$P\{X(t_2)=j\mid X(t'),0\leqslant t'\leqslant t_1\}=P\{X(t_2)=j\mid X(t_1)=i\},t_1\leqslant t_2,i,j\in S$$

记

$$p_{ij}(t_1,t_2)\triangleq P\{X(t_2)=j\mid X(t_1)=i\}$$

称此条件概率为连续时间的 Markov 链的转移概率.

显然有

$$\begin{cases}p_{ij}(t_1,t_2)\geqslant 0\\ \sum\limits_{j\in I}p_{ij}(t_1,t_2)=1,i\in I\end{cases}$$

如果 $p_{ij}(t_1,t_2)$ 仅为时间差 $t=t_2-t_1$ 的函数，而与 t_1 和 t_2 的值无关，则称此连续时间的 Markov 链为齐次的或者时齐的.

定理 5.1　连续时间的齐次 Markov 链的转移概率 $p_{ij}(t)$ 满足如下性质：

$$p_{ij}(t)=p_{ij}(t_1,t_2)\triangleq P\{X(t_2)=j\mid X(t_1)=i\},t=t_2-t_1$$

(1) $p_{ij}(t)\geqslant 0,i,j\in I,t\geqslant 0$；

(2) $\sum\limits_{j\in I}p_{ij}(t)=1,i\in I,t\geqslant 0$；

(3) 连续时间的齐次 Markov 链的 C-K 方程为

$$p_{ij}(t+\tau)=\sum_{k\in I}p_{ik}(t)p_{kj}(\tau)\quad(i,j\in I,t>0,\tau>0)$$

证明　由转移概率的定义可知,性质(1)和(2)显然成立.下面只证(3):由全概率公式及马尔可夫性得

$$\begin{aligned}
&p_{ij}(t+\tau)=P\{X(t+\tau)=j\mid X(0)=i\}\\
&=\sum_{k\in I}P\{X(t+\tau)=j,X(t)=k\mid X(0)=i\}\\
&=\sum_{k\in I}P\{X(t)=k\mid X(0)=i\}P\{X(t+\tau)=j\mid X(t)=k\}\\
&=\sum_{k\in I}P\{X(t)=k\mid X(0)=i\}P\{X(\tau)=j\mid X(0)=k\}\\
&=\sum_{k\in I}p_{ik}(t)p_{kj}(\tau)
\end{aligned}$$

定义 5.2　称一个连续时间的 Markov 链是正则的,若它以概率1在任意有限长的时间内转移的次数是有限的.

一般地,对于转移概率 $p_{ij}(t)$,假定它满足如下的正则条件(或连续性条件):

$$\lim_{t\to 0}p_{ij}(t)=\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$$

本节恒假定 Markov 链满足正则条件.正则条件说明,过程刚进入某个状态不可能立即又跳跃到另一个状态.这正好说明一个物理系统要在有限时间内发生无限多次跳跃,从而消耗无穷多的能量是不可能的.

下面主要讨论连续时间的齐次 Markov 链.

定理 5.2　设 $\{X(t),t\geqslant 0\}$ 是连续时间的 Markov 链,假定时刻0过程刚刚到达 $i(i\in S)$.以 τ_i 记过程在离开 i 之前在 i 停留的时间,则 τ_i 服从指数分布.

证明　只需证明对 $s,t\geqslant 0$,有

$$P\{\tau_i>s+t\mid \tau_i>s\}=P\{\tau_i>t\}$$

即无记忆性.

注意到

$$\{\tau_i>s\}\Leftrightarrow\{X(u)=i,0\leqslant u\leqslant s\mid X(0)=i\}$$

$$\{\tau_i>s+t\}\Leftrightarrow\{X(u)=i,0\leqslant u\leqslant s+t\mid X(0)=i\}$$

则

$$\begin{aligned}
&P\{\tau_i>s+t\mid \tau_i>s\}\\
&=P\{X(u)=i,0\leqslant u\leqslant s+t\mid X(u)=i,0\leqslant u\leqslant s\}\\
&=P\{X(u)=i,s\leqslant u\leqslant s+t\mid X(s)=i\}\\
&=P\{X(u)=i,0\leqslant u\leqslant t\mid X(0)=i\}\\
&=P\{\tau_i>t\}
\end{aligned}$$

其中第三个等式利用马氏性,第四个等式利用时齐性.

上述定理说明,连续时间的 Markov 链具有如下两条性质:

(1) 在转移到下一个状态之前，处于状态 i 的时间服从参数为 μ_i 的指数分布.

(2) 在过程离开状态 i 时，将以概率 p_{ij} 到达 j，且 $\sum_{j\in S} p_{ij} = 1$.

定义 5.3 若 $\mu_i = +\infty$，则它在状态 i 停留的平均时间为 0，即一旦进入马上离开，称 i 是瞬过的. 若对任意 $i, 0 \leqslant \mu_i < +\infty$，则称 Markov 链不存在瞬过状态.

【例 5.1】 试证明泊松过程 $\{X(t), t \geqslant 0\}$ 为连续时间的齐次 Markov 链.

证明 先证泊松过程具有马尔可夫性，再证齐次性.

由于它是独立增量过程，且 $X(0) = 0$，对任意 $0 < t_1 < t_2 < \cdots < t_n < t_{n+1}$ 有

$$\begin{aligned} & P\{X(t_{n+1}) = i_{n+1} \mid X(t_1) = i_1, X(t_2) = i_2, \cdots, X(t_n) = i_n\} \\ = & P\{X(t_{n+1}) - X(t_n) = i_{n+1} - i_n \mid X(t_1) - X(0) = i_1, \cdots, X(t_n) - X(t_{n-1}) = i_n - i_{n-1}\} \\ = & P\{X(t_{n+1}) - X(t_n) = i_{n+1} - i_n\} \end{aligned}$$

又因为

$$\begin{aligned} & P\{X(t_{n+1}) = i_{n+1} \mid X(t_n) = i_n\} \\ = & P\{X(t_{n+1}) - X(t_n) = i_{n+1} - i_n \mid X(t_n) - X(0) = i_n\} \\ = & P\{X(t_{n+1}) - X(t_n) = i_{n+1} - i_n\} \end{aligned}$$

所以泊松过程具有马尔可夫性.

下面再证齐次性.

当 $j \geqslant i$ 时，由泊松过程的定义得

$$P\{X(s+t) = j \mid X(s) = i\} = P\{X(s+t) - X(s) = j - i\} = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}$$

当 $j < i$ 时，$p_{ij}(s,t) = p_{ij}(t) = 0$.

即转移概率只与 t 有关，泊松过程具有齐次性.

定义 5.4 对齐次马尔可夫过程在开始时刻 $(t = 0)$ 取状态 j 的概率分布

$$p_j(0) = P\{X(0) = j\}, j \in I$$

称为它的初始分布.

对齐次马尔可夫过程在 $t (t \geqslant 0)$ 时刻取状态 j 的概率分布

$$p_j(t) = P\{X(t) = j\}, j \in I$$

称为它在 t 时刻的绝对概率分布.

不加证明地给出绝对概率分布的以下性质：

(1) $p_j(t) \geqslant 0$；

(2) $\sum_{j\in I} p_j(t) = 1$；

(3) $p_j(t) = \sum_{i\in I} p_i(0) p_{ij}(t)$；

(4) $p_j(t+\tau) = \sum_{i\in I} p_i(t) p_{ij}(\tau)$；

(5) $P\{X(t_1)=i_1, X(t_2)=i_2, \cdots, X(t_n)=i_n\}$

$$= \sum_{i\in I} p_{i_0}(0) p_{i_0 i_1}(t_1) p_{i_1 i_2}(t_2-t_1)\cdots p_{i_{n-1}i_n}(t_n-t_{n-1}).$$

5.2　柯尔莫哥洛夫-费勒(Kolmogrov-Feller)微分方程

定理5.3　设齐次马尔可夫过程满足正则条件，则对任意固定的 $i,j\in I=\{0,1,2,\cdots\}$，$p_{ij}(t)$ 是 t 的一致连续函数.

证明　设 $h>0$，则

$$\begin{aligned} p_{ij}(t+h)-p_{ij}(t) &= \sum_{k=0}^{+\infty} p_{ik}(h)p_{kj}(t)-p_{ij}(t) \\ &= p_{ii}(h)p_{ij}(t)-p_{ij}(t)+\sum_{k\neq i}^{+\infty} p_{ik}(h)p_{kj}(t) \\ &= -[1-p_{ii}(h)]p_{ij}(t)+\sum_{k\neq i}^{+\infty} p_{ik}(h)p_{kj}(t) \end{aligned}$$

从而得到

$$p_{ij}(t+h)-p_{ij}(t) \geqslant -[1-p_{ii}(h)]p_{ij}(t) \geqslant -[1-p_{ii}(h)]$$

和

$$p_{ij}(t+h)-p_{ij}(t) \leqslant \sum_{k\neq i}^{+\infty} p_{ik}(h)p_{kj}(t) \leqslant \sum_{k\neq i}^{+\infty} p_{ik}(h) = 1-p_{ii}(h)$$

因此得到

$$|p_{ij}(t+h)-p_{ij}(t)| \leqslant 1-p_{ii}(h)$$

对于 $h<0$，可得类似的不等式，所以有

$$|p_{ij}(t+h)-p_{ij}(t)| \leqslant 1-p_{ii}(|h|) \to 0 \quad (h\to 0)$$

即 $p_{ij}(t)$ 是 t 的一致连续函数.

定理5.4　设 $p_{ij}(t)$ 是齐次马尔可夫过程的转移概率，则 $P_{ij}(t)$ 具有下列极限性质：

(1) $\lim\limits_{t\to 0} \dfrac{1-p_{ii}(t)}{t} = q_{ii} \leqslant +\infty, i=j$;

(2) $\lim\limits_{t\to 0} \dfrac{p_{ij}(t)}{t} = q_{ij} < +\infty, \quad i\neq j$.

称 q_{ij} 为从状态 i 到 j 的转移速率或者跳跃强度.

证明　参见文献[13].

定理5.4中极限的概率意义表明，在长度为 t 的时间区间内，过程从状态 i 转移

到另一其他状态的转移概率 $1-p_{ii}(t)$ 等于 q_{ii} 加上一个比 t 高阶的无穷小量;而从状态 i 转移到状态 j 的概率 $p_{ij}(t)$ 等于 q_{ij} 加上一个比 t 高阶的无穷小量.

推论 5.1 对有限状态时齐的连续时间 Markov 链,有

$$q_{ii}=\sum_{j\neq i}q_{ij}<+\infty$$

证明 由定理 5.1 知,$\sum_{j\in I}p_{ij}(t)=1$,即

$$1-p_{ii}(t)=\sum_{j\neq i}p_{ij}(t)$$

故

$$\lim_{t\to 0}\frac{1-p_{ii}(t)}{t}=\lim_{t\to 0}\sum_{j\neq i}\frac{p_{ij}(t)}{t}=\sum_{j\neq i}\lim_{t\to 0}\frac{p_{ij}(t)}{t}=\sum_{j\neq i}q_{ij}<+\infty$$

注:对于无限状态的情况,一般只能得到 $q_{ii}\geqslant\sum_{j\neq i}q_{ij}$. 为了简单起见,设状态空间 I 为$\{0,1,2,\cdots\}$,此时记

$$\boldsymbol{Q}=\begin{pmatrix} q_{00} & q_{01} & q_{02} & \cdots & q_{0n} \\ q_{10} & q_{11} & q_{12} & \cdots & q_{1n} \\ \vdots & \vdots & \vdots & & \vdots \\ q_{n0} & q_{n1} & q_{n2} & \cdots & q_{nn} \end{pmatrix}_{(n+1)\times(n+1)}$$

称之为齐次马尔可夫过程的转移速率矩阵,简称 $\boldsymbol{Q}$ 矩阵. 当矩阵元素 $q_{ii}=\sum_{j\neq i}q_{ij}<+\infty$ 时,则称该矩阵是保守的.

由于对 $i\in I$,$p_{ij}(t)$ 在 $t=0$ 处的导数(右导数)存在,即 $\lim_{t\to 0}\frac{p_{ii}(t)-\delta_{ij}}{t}=q_{ij}=p'_{ij}(0^+)$,所以 $\boldsymbol{Q}$ 矩阵也称为转移概率矩阵 $P(t)$ 的密度矩阵. 以 $\boldsymbol{Q}$ 为密度矩阵的马尔可夫过程称为 $\boldsymbol{Q}$ 过程. 由 $\boldsymbol{Q}$ 矩阵求转移概率矩阵 $\boldsymbol{P}(t)$ 的问题,称为 $\boldsymbol{Q}$ 矩阵问题.

$\boldsymbol{Q}$ 矩阵反映了齐次马尔可夫过程的转移概率(函数)在 $t=0$ 处的变化率,它是一个常数矩阵,而转移概率矩阵是一个函数矩阵,在计算上,$\boldsymbol{Q}$ 矩阵较转移概率矩阵易于获得,因而给解决实际问题带来了较大的方便.

转移速率 q_{ij} 具有如下性质:

(1)$q_{ij}\leqslant 0,i=j,i,j\in I$;

(2)$q_{ij}\geqslant 0,i\neq j,i,j\in I$;

(3)$\sum_{j\in I}q_{ij}=0$.

对任意 $i\in I$,下面的极限

$$-p'_{ii}(0)=\lim_{t\to 0^+}\frac{1-p_{ii}(t)}{t}=q_{ii}\leqslant+\infty$$

存在,即导数存在,但可能为无穷大. q_{ii} 的取值决定了状态的性质.

(1) 当 $q_{ii}=0$ 时,$p_{ii}(t)=1(\forall t>0)$,表明系统从 i 出发不发生转移,称状态

i 为**吸收状态**.

(2) 当 $q_{ii}=+\infty$ 时,有 $p_{ii}(t)=0(\forall t>0)$,表明系统从 i 出发经任意时间都不回到 i,称状态 i 为**瞬时状态**.

(3) 当 $0<q_{ii}<+\infty$ 时,$0<p_{ii}(t)<1(\forall t>0)$,称状态 i 为**逗留状态**.

若令 $\tau(\omega)$ 表示系统一直停留在出发状态的时间长度,则 $E[\tau(\omega)\mid X(0)=i]=1/q_{ii}$,即 $1/q_{ii}$ 表示停留在状态 i 的平均时间长度.

定理 5.5 (Kolmogrov-Feller 微分方程) 对一切 $i,j\in I,t\geqslant 0$ 且 $\sum\limits_{j\neq i}q_{ij}=q_{ii}<+\infty$ 有

(1) 向后方程

$$p'_{ij}(t)=\sum_{k\neq i}q_{ik}p_{kj}(t)-q_{ii}p_{ij}(t)$$

用矩阵表示为 $\boldsymbol{P}'(t)=\boldsymbol{QP}(t)$.

(2) 在适当的正则条件下,有向前方程

$$p'_{ij}(t)=\sum_{k\neq j}q_{kj}p_{ik}(t)-q_{jj}p_{ij}(t)$$

用矩阵表示为 $\boldsymbol{P}'(t)=\boldsymbol{P}(t)\boldsymbol{Q}$.

证明 只证有限状态情形.

由 C-K 方程有

$$p_{ij}(t+h)=\sum_{k\in I}p_{ik}(h)p_{kj}(t)$$

或等价地

$$p_{ij}(t+h)-p_{ii}(h)p_{ij}(t)=\sum_{k\neq i}p_{ik}(h)p_{kj}(t)$$

变形为

$$p_{ij}(t+h)-p_{ij}(t)=\sum_{k\neq i}p_{ik}(h)p_{kj}(t)-[1-p_{ii}(h)]p_{ij}(t)$$

于是

$$\lim_{h\to 0}\frac{p_{ij}(t+h)-p_{ij}(t)}{h}=\lim_{h\to 0}\sum_{k\neq i}\frac{p_{ik}(h)}{h}p_{kj}(t)-\lim_{h\to 0}\frac{1-p_{ii}(h)}{h}p_{ij}(t)$$

即证 $p'_{ij}(t)=\sum\limits_{k\neq i}q_{ik}p_{kj}(t)-q_{ii}p_{ij}(t)$.

利用 $p_{ij}(t+h)=\sum\limits_{k\in I}p_{ik}(t)p_{kj}(h)$ 同理可证向前方程.

考虑状态有限时,若令

$$\boldsymbol{\Gamma}_i(t)=(p_{i0}(t),p_{i1}(t),\cdots,p_{in}(t))$$

则有

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{\Gamma}_i(t)}{\mathrm{d}t}=\boldsymbol{\Gamma}_i(t)\boldsymbol{Q},i=0,1,2,\cdots,n\\ \boldsymbol{\Gamma}_i(0)=(0,0,\cdots,1,\cdots,0)\end{cases}$$

进一步,若记

$$\boldsymbol{P}(t)=\begin{pmatrix}\boldsymbol{\Gamma}_0(t)\\ \boldsymbol{\Gamma}_1(t)\\ \vdots\\ \boldsymbol{\Gamma}_n(t)\end{pmatrix}=\begin{pmatrix}p_{00}(t) & p_{01}(t) & \cdots & p_{0n}(t)\\ p_{10}(t) & p_{11}(t) & \cdots & p_{1n}(t)\\ \vdots & \vdots & & \vdots\\ p_{n0}(t) & p_{n1}(t) & \cdots & p_{nn}(t)\end{pmatrix}_{(n+1)\times(n+1)}$$

则有

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{P}(t)}{\mathrm{d}t}=\boldsymbol{P}(t)\boldsymbol{Q}\\ \boldsymbol{P}(0)=\boldsymbol{I}_{(n+1)\times(n+1)}\end{cases}$$

此即为 Kolmogrov-Feller 向前方程的矩阵形式.

若记

$$\boldsymbol{S}_j(t)=\begin{pmatrix}p_{0j}(t)\\ p_{1j}(t)\\ \vdots\\ p_{nj}(t)\end{pmatrix}$$

则有

$$\frac{\mathrm{d}\boldsymbol{S}_j(t)}{\mathrm{d}t}=\boldsymbol{Q}\boldsymbol{S}_j(t),j=0,1,2,\cdots,n$$

初始条件为

$$\boldsymbol{S}_j(0)=\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix}(j+1)$$

即

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{P}(t)}{\mathrm{d}t}=\boldsymbol{Q}\boldsymbol{P}(t)\\ \boldsymbol{P}(0)=\boldsymbol{I}_{(n+1)\times(n+1)}\end{cases}$$

上面的方程组即为 Kolmogrov-Feller 向后方程的矩阵形式.

Kolmogrov-Feller 微分方程主要是建立了齐次马尔可夫过程的 $\boldsymbol{Q}$ 矩阵与转移概率矩阵 $\boldsymbol{P}(t)$ 之间的关系. 若给定了密度矩阵 $\boldsymbol{Q}$,在一定的条件下,可以通过 Kolmogrov-Feller 微分方程解出转移概率 $p_{ij}(t)$. 例如,如果 $\boldsymbol{Q}$ 是一个有限维矩阵,则由向后与向前方程可以求解转移概率矩阵 $\boldsymbol{P}(t)$ 为

$$P(t)=e^{Qt}=\sum_{j=0}^{+\infty}[(Qt)^j/j!]$$

定理5.6　齐次马尔可夫过程在 t 时刻处于状态 $j\in I$ 的绝对概率 $p_j(t)$ 满足下列方程：

$$p'_j(t)=-q_{jj}p_j(t)+\sum_{k\neq j}q_{kj}p_k(t)$$

证明　由齐次马尔可夫过程绝对概率分布的性质有

$$p_j(t)=\sum_{i\in I}p_i(0)p_{ij}(t)$$

将 Kolmogrov-Feller 向前微分方程两边同乘以 $p_i(0)$，并对 i 求和得

$$\sum_{i\in I}p_j(0)p'_{ij}(t)=\sum_{i\in I}\sum_{k\neq j}p_j(0)q_{kj}p_{ik}(t)-\sum_{i\in I}p_j(0)q_{jj}p_{ij}(t)$$

于是有

$$p'_j(t)=-q_{jj}p_j(t)+\sum_{k\neq j}q_{kj}p_k(t)$$

与第4章讨论离散马尔可夫链类似，不加证明地给出连续时间马尔可夫链的转移概率 $p_{ij}(t)(t\to+\infty)$ 的极限分布与平稳分布的有关性质.

定义5.5　设 $p_{ij}(t)$ 为连续时间马尔可夫链的转移概率，若存在时刻 t_1 和 t_2，使得 $p_{ij}(t_1)>0,p_{ij}(t_2)>0$，则称状态 i 和 j 是**互通**的. 若所有状态都是互通的，则称此马尔可夫链为**不可约的**.

定理5.7　取固定的 $h>0$，$p_{ij}(h)$ 为一步转移概率，则 n 步转移概率为 $p_{ij}^{(n)}(t)=p_{ij}(nh)$.

(1) 设 $\{p_{ij}(h)\}$ 是非周期不可约的，且对任意 $i,j\in I$，下列极限存在且与 i 无关：

$$\lim_{n\to+\infty}p_{ij}^{(n)}(h)=\lim_{n\to+\infty}p_{ij}(nh)=\pi_j(h)$$

即

$$\lim_{n\to+\infty}p_{ij}^{(n)}(h)=\begin{cases}0, & j\text{ 为非常返或零常返态}\\ \dfrac{1}{\mu_j(h)}, & j\text{ 为遍历态}\end{cases}$$

(2)(马尔可夫定理) 对一切 $i,j\in I$，下列极限存在且与 i 无关：

$$\lim_{t\to+\infty}p_{ij}(t)=\pi_j$$

(3) $\{\pi_j,j\in I\}$ 有下列性质：$\pi_j\geqslant 0$，$\sum_{j\in I}\pi_j=1$，对任意 $t>0$ 有

$$\pi_j=\sum_{k\in I}\pi_k p_{kj}(t)$$

(4) 若对某一 $j\in I$ 有 $\pi_j>0$，则对一切 $i\in I$，有 $\pi_i>0$，且有

$$\sum_{j\in I}\pi_j = 1, \sum_{i\neq j}\pi_i q_{ij} = \pi_j q_j, j \in I$$

定理 5.8 设连续时间马尔可夫链是不可约的,则存在下列性质:

(1) 若 I 是正常返态的,则 $\lim\limits_{t\to+\infty} p_{ij}(t)$ 存在且等于 π_j. $\pi_j > 0, j \in I$,其中 π_j 是方程组

$$\begin{cases}\sum\limits_{k\neq j}\pi_k q_{kj} = \pi_j q_{jj} \\ \sum\limits_{j\in I}\pi_j = 1\end{cases}$$

的唯一非负解,此时称$\{\pi_j, j \in I\}$是该过程的平稳分布,且有

$$\lim_{t\to+\infty} p_j(t) = \pi_j$$

(2) 若它是零常返态的或非常返态的,则

$$\lim_{t\to+\infty} p_{ij}(t) = \lim_{t\to+\infty} p_j(t) = 0, \quad i,j \in I$$

【例 5.2】 (随机信号问题)计算机中某个触发器有两种状态,记为"0"和"1".设触发器状态的变化构成一个齐次马尔可夫过程$\{X(t), t \in [0, +\infty)\}$, $I = \{0, 1\}$,且有

$$p_{01}(\Delta t) = \lambda\Delta t + o(\Delta t), p_{10}(\Delta t) = \mu\Delta t + o(\Delta t)$$

求矩阵 $\boldsymbol{Q}$ 与 $\boldsymbol{P}(t)$.

解 设 $X(t)$ 表示时刻 t 触发器的状态,则

$$q_{01} = \lim_{\Delta t\to 0^+}\frac{p_{01}(\Delta t) - \delta_{01}}{\Delta t} = \lambda$$

$$q_{10} = \lim_{\Delta t\to 0^+}\frac{p_{10}(\Delta t) - \delta_{10}}{\Delta t} = \mu$$

$$q_{00} = -q_{01} = -\lambda, q_{11} = -q_{10} = -\mu$$

于是 $\boldsymbol{Q} = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$, $\boldsymbol{P}(t) = \begin{bmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{bmatrix}$, $\boldsymbol{P}(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

由 Kolmogrov-Feller 向前方程,可得

$$\frac{\mathrm{d}p_{00}(t)}{\mathrm{d}t} = -\lambda p_{00}(t) + \mu p_{01}(t)$$

$$\frac{\mathrm{d}p_{01}(t)}{\mathrm{d}t} = \lambda p_{00}(t) - \mu p_{01}(t)$$

$$\frac{\mathrm{d}p_{10}(t)}{\mathrm{d}t} = -\lambda p_{10}(t) + \mu p_{11}(t)$$

$$\frac{\mathrm{d}p_{11}(t)}{\mathrm{d}t} = \lambda p_{10}(t) - \mu p_{11}(t)$$

$$p_{00}(0) = p_{11}(0) = 1, p_{10}(0) = p_{01}(0) = 0$$

解上面的线性微分方程组,得

$$p_{00}(t)=\mu_0+\lambda_0 e^{-(\lambda+\mu)t},\ p_{01}(t)=\lambda_0[1-e^{-(\lambda+\mu)t}]$$
$$p_{10}(t)=\mu_0[1-e^{-(\lambda+\mu)t}],\ p_{11}(t)=\lambda_0+\mu_0 e^{-(\lambda+\mu)t}$$

其中 $\lambda_0=\dfrac{\lambda}{\lambda+\mu},\mu_0=\dfrac{\mu}{\lambda+\mu}$.

【例5.3】（随机游动问题）质点在线段[1,5]上作随机游动,且只能停留在整数点上.质点可在任何时刻发生移动,移动规则是:

① 若时刻 t 质点位于2,3,4处,则在 $(t,t+\Delta t)$ 内右移一格的概率是 $\lambda\Delta t+o(\Delta t)$,左移一格的概率是 $\mu\Delta t+o(\Delta t)$,停留在原处的概率是 $1-\lambda\Delta t-\mu\Delta t+o(\Delta t)$;

② 若时刻 t 质点位于1处,则在 $(t,t+\Delta t)$ 内右移一格的概率是 $\lambda\Delta t+o(\Delta t)$,停留在原处的概率是 $1-\lambda\Delta t+o(\Delta t)$;

③ 若时刻 t 质点位于5处,则在 $(t,t+\Delta t)$ 内左移一格的概率是 $\mu\Delta t+o(\Delta t)$,停留在原处的概率是 $1-\mu\Delta t+o(\Delta t)$;

④ 在 $(t,t+\Delta t)$ 内发生其他移动的概率都是 $o(\Delta t)$.

求各 $p_{ij}(t)$ 满足的微分方程组.

解　按移动规则写出转移概率函数

$$p_{ij}(\Delta t)=\begin{cases}\lambda\Delta t+o(\Delta t), & j=i+1,\\ \mu\Delta t+o(\Delta t), & j=i-1,\\ 1-\lambda\Delta t-\mu\Delta t+o(\Delta t), & j=i=2,3,4\\ o(\Delta t), & \text{其他}.\end{cases}$$

由极限式可求得

$$q_{ij}=\begin{cases}\lambda, & j=i+1,\\ \mu, & j=i-1,\\ 1-\lambda-\mu, & j=i=2,3,4,\\ 0, & \text{其他}.\end{cases}$$

类似地,有 $p_{1j}(\Delta t)=\begin{cases}\lambda\Delta t+o(\Delta t), & j=2,\\ 1-\lambda\Delta t+o(\Delta t), & j=1,\\ o(\Delta t), & \text{其他},\end{cases}\quad q_{1j}=\begin{cases}\lambda, & j=2,\\ -\lambda, & j=1,\\ 0, & \text{其他};\end{cases}$

$$p_{5j}(\Delta t)=\begin{cases}\mu\Delta t+o(\Delta t), & j=4,\\ 1-\mu\Delta t+o(\Delta t), & j=5,\\ o(\Delta t), & \text{其他},\end{cases}\quad q_{5j}=\begin{cases}\mu, & j=4,\\ -\mu, & j=5,\\ 0, & \text{其他}.\end{cases}$$

综合上述结果,得

$$Q=\begin{bmatrix}-\lambda & \lambda & & & \\ \mu & -(\lambda+\mu) & \lambda & & \\ & \mu & -(\lambda+\mu) & \lambda & \\ & & \mu & -(\lambda+\mu) & \lambda \\ & & & \mu & -\mu\end{bmatrix}$$

故 $\frac{\mathrm{d}\boldsymbol{P}(t)}{\mathrm{d}t}=\boldsymbol{P}(t)\boldsymbol{Q},\frac{\mathrm{d}\boldsymbol{P}(t)}{\mathrm{d}t}=\boldsymbol{Q}\boldsymbol{P}(t)$.

$$\boldsymbol{P}(0)=\begin{pmatrix}1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{pmatrix}$$

解方程组即得转移概率函数 $p_{ij}(t)$.

【例 5.4】 由泊松过程定义知

$$p_{i,i+1}(\Delta t)=P\{N(t+\Delta t)-N(t)=1\mid N(t)=i\}=\lambda\Delta t+o(\Delta t)$$
$$p_{ii}(\Delta t)=P\{N(t+\Delta t)-N(t)=0\mid N(t)=i\}=1-\lambda\Delta t+o(\Delta t)$$

求 Kolmogrov-Feller 微分方程的解.

解 因为

$$\lim_{\Delta t\to 0}\frac{1-p_{ii}(\Delta t)}{\Delta t}=q_{ii}=\lambda,\lim_{\Delta t\to 0}\frac{p_{i,i+1}(\Delta t)}{\Delta t}=q_{i,i+1}=\lambda$$

所以 $p'_{ij}(t)=q_{i,i+1}p_{i+1,j}(t)-q_{ii}p_{ij}(t)=\lambda p_{i+1,j}(t)-\lambda p_{ij}(t)$.

当 $j=i$ 时,有 $p'_{ii}(t)=-\lambda p_{ii}(t)$;

当 $j=i+1$ 时,有 $p'_{i,i+1}(t)=\lambda p_{i+1,i+1}(t)-\lambda p_{i,i+1}(t)$;

当 $j=i+2$ 时,有 $p'_{i,i+2}(t)=\lambda p_{i+1,i+2}(t)$;

在其他情形下,微分方程不存在.

由条件 $p_{ii}(0)=1$ 得微分方程的解为

$$p_{ii}(t)=\mathrm{e}^{-\lambda t},p_{i,i+1}(t)=\lambda t\mathrm{e}^{-\lambda t}$$
$$p_{ij}(t)=\mathrm{e}^{-\lambda t}\frac{(\lambda t)^{j-i}}{(j-i)!},\quad j\geqslant i\geqslant 0$$

【例 5.5】 设有参数连续、状态离散的马氏过程$\{X(t),t\geqslant 0\}$,状态空间为 $I=\{1,2,\cdots,m\}$,当 $i\neq j,i,j=1,2,\cdots,m$ 时,$q_{ij}=1,q_{ii}=-(m-1),i=1,2,\cdots,m$,求 $p_{ij}(t)$.

解 由 Kolmogrov-Feller 向前方程得

$$\frac{\mathrm{d}p_{ij}(t)}{\mathrm{d}t}=-(m-1)p_{ij}(t)+\sum_{k\neq j,k\in I}p_{ik}(t)$$

由

$$\sum_{k\in I}p_{ik}(t)=1$$

可知

$$\sum_{k \neq j, k \in I} p_{ik}(t) = 1 - p_{ij}(t)$$

因此，有

$$\frac{\mathrm{d}p_{ij}(t)}{\mathrm{d}t} = -(m-1)p_{ij}(t) + [1 - p_{ij}(t)] = 1 - mp_{ij}(t) \quad (i,j = 1,2,\cdots,m)$$

解此微分方程，得

$$p_{ij}(t) = ce^{mt} + \frac{1}{m} \quad (i,j = 1,2,\cdots,m)$$

利用初始条件

$$p_{ii}(0) = 1, p_{ij}(0) = 0 \quad (i \neq j)$$

可得

$$p_{ii}(t) = \left(1 - \frac{1}{m}\right)\mathrm{e}^{-mt} + \frac{1}{m} \quad (i = 1,2,\cdots,m)$$

$$p_{ij}(t) = \frac{1}{m}(1 - \mathrm{e}^{-mt}) \quad (i \neq j, i,j = 1,2,\cdots,m)$$

5.3　生灭过程

本节主要介绍在实际问题中使用较为广泛的一类特殊的连续时间马尔可夫链模型，即生灭过程. 具体定义如下：

定义 5.6　一个状态空间为 $I = \{0,1,2,\cdots\}$、有连续参数集 $T = [0,+\infty)$ 的齐次马尔可夫过程 $\{X(t), t \geq 0\}$，如果它的密度矩阵为

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & & & \\ \mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & & \\ & \mu_2 & -(\mu_2+\lambda_2) & \lambda_2 & \\ & & \mu_3 & -(\mu_3+\lambda_3) & \lambda_3 \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

则称该过程为**生灭过程**，其中 λ_i 为**出生率**，μ_i 为**死亡率**；Q 必为保守阵.

生灭过程的转移概率 $p_{ij}(t)$ 的性质：设 $i \in I$，对充分小的时间 $t > 0$，有

$$p_{i,i+1}(t) = \lambda_i t + o(t), \quad \lambda_i > 0$$

$$p_{i,i-1}(t) = \mu_i t + o(t), \quad \mu_0 = 0, \mu_i > 0, i \geq 1$$

$$p_{ii}(t) = 1 - (\lambda_i + \mu_i)t + o(t)$$

$$p_{ij}(t) = o(t), \quad |i - j| \geq 2$$

若 $\lambda_i = i\lambda$，$\mu_i = i\mu$（λ, μ 为正常数），则称 $\{X(t), t \geq 0\}$ 为**线性生灭过程**.

若在密度矩阵Q中,$\mu_i = 0(i \in I)$,则群体个体减少的概率为零,称为**纯生过程**(如泊松过程).若在密度矩阵Q中,$\lambda_i = 0(i \in I)$,则群体个数增加的概率为零,称为**纯灭过程**.

生灭过程的概率意义可作如下解释:由生灭过程的转移概率$p_{ij}(t)$的性质可知,如果忽略t的高阶无穷小量,生灭过程即种群群体的状态变化有以下三种可能:

(1) 由状态$i \to i+1$,即增加了一个个体,概率是$\lambda_i t$;

(2) 由状态$i \to i-1$,即减少了一个个体,概率是$\mu_i t$;

(3)$i \to i$,即群体个数无变化.

由此可以得出,生灭过程的所有状态是互通的,但在充分小的时间内,只能在两个相邻状态内变化,或者状态无变化.

类似于5.2节的推导过程,可得到生灭过程的柯尔莫哥洛夫-费勒微分方程和平稳分布的相关结果如下:

1. 生灭过程的柯尔莫哥洛夫-费勒微分方程

(1) 向后方程是

$$\begin{cases} p'_{ij}(t) = -(\mu_i + \lambda_i)p_{ij}(t) + \lambda_i p_{i+1,j}(t) + \mu_i p_{i-1,j}(t), i \geqslant 1, j \in I \\ p'_{0j}(t) = -\lambda_0 p_{0j}(t) + \lambda_0 p_{1j}(t), j \in I \end{cases}$$

(2) 向前方程是

$$\begin{cases} p'_{ij}(t) = -p_{ij}(t)(\mu_j + \lambda_j) + p_{i,j-1}(t)\lambda_{j-1} + p_{i,j+1}(t)\mu_{j+1}, j \geqslant 1, i \in I \\ p'_{i0}(t) = -p_{i0}(t)\lambda_0 + p_{i1}(t)\mu_1, i \in I \end{cases}$$

绝对概率有如下方程:

$$\begin{cases} p'_j(t) = -p_j(t)(\mu_j + \lambda_j) + p_{j-1}(t)\lambda_{j-1} + p_{j+1}(t)\lambda_{j+1}, j \geqslant 2 \\ p'_0(t) = -p_0(t)\lambda_0 + p_1(t)\mu_1 \end{cases}$$

2. 生灭过程的平稳分布

$$\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_n, \cdots), \pi_0 = \left(1 + \sum_{n=1}^{+\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n}\right)^{-1}$$

$$\pi_n = \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n}\pi_0, n = 1,2,\cdots$$

在级数$1 + \sum\limits_{n=1}^{+\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n}$收敛的条件下,以$Q$为密度矩阵的生灭过程存在平稳分布.由$\lim\limits_{t \to +\infty} p_{ij}(t) = \pi_j, i,j \in I$知,当$t$充分大时,$p_{ij}(t) \approx \pi_j$.进一步得到生灭过程存在唯一平稳分布的充要条件:

定理5.9 设$\{X(t), t \geqslant 0\}$为一生灭过程,其中$\lambda_i > 0(i \geqslant 0)$;$\mu_i > 0(i \geqslant 1)$,$\mu_0 = 0$,则$\{X(t), t \geqslant 0\}$存在唯一的平稳分布(它就等于极限分布)的充要条件为

$$\sum_{n=1}^{+\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n} < +\infty$$

【例5.6】（电话问题的爱尔朗(Erlang)公式）某交换台有s条中继线，区内用户与区外用户通话要通过中继线．由于用户数量很大，可以认为不管占用了几条中继线，不通话的用户数量可看作是不变的，因此，假定在$(t,t+\Delta t)$内又有一用户要同区外通话的概率是$\lambda\Delta t+o(\Delta t)$，与正在通话的户数无关．若中继线有空则接通，否则用户的要求被取消．而在时刻t正与区外通话的用户能在$(t,t+\Delta t)$内结束通话，从而空出一条中继线的概率是$\mu\Delta t+o(\Delta t)$．设各用户与区外通话是独立的，若用$X(t)$表示时刻t正在使用的中继线个数，则$\{X(t),t\geqslant 0\}$是齐次马尔可夫过程，求其平稳分布并导出爱尔朗公式．

解　转移概率为

$$\begin{cases} p_{i,i+1}(\Delta t)=\lambda\Delta t+o(\Delta t), & i=0,1,\cdots,s-1 \\ p_{i,i-1}(\Delta t)=\mu i\Delta t+o(\Delta t), & i=1,2,\cdots,s \\ p_{ii}(\Delta t)=1-(\lambda+\mu i)\Delta t+o(\Delta t), & i=0,1,\cdots,s-1 \\ p_{ss}(\Delta t)=1-s\mu\Delta t+o(\Delta t) \\ p_{ij}(\Delta t)=0,\ |i-j|>1 \end{cases}$$

显然，$X(t)$是一个生灭过程．

$$\lambda_i=\lambda,\quad i=0,1,\cdots,s-1$$

$$\mu_i=i\mu,\quad i=1,2,\cdots,s$$

其平稳分布$\boldsymbol{\pi}=(\pi_0,\pi_1,\cdots,\pi_s)$为

$$\pi_k=\frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k}\pi_0=\frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k\pi_0,\quad k=1,2,\cdots,s$$

$$\pi_0=\left[1+\sum_{k=1}^{s}\frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k\right]^{-1}=\left[\sum_{k=0}^{s}\frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k\right]^{-1}$$

于是，得到爱尔朗公式

$$\pi_k=\frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k\bigg/\sum_{l=0}^{s}\frac{1}{l!}\left(\frac{\lambda}{\mu}\right)^l,\quad k=1,2,\cdots,s$$

【例5.7】（酶的产生过程）设有大量的实验样品，每个样品包含一个酶分子和母体物质．若在$(t,t+\Delta t)$内形成第二个酶分子的样品数与时刻t未形成第二个酶分子的样品数成正比，且与时刻t无关．任一包含一个酶分子的样品在$(t,t+\Delta t)$内产生第二个酶分子的概率为$\lambda\Delta t+o(\Delta t)$．设在时间间隔$\Delta t$内增加两个或更多酶分子的概率是$o(\Delta t)$．令$X(t)$表示时刻$t$系统中的酶分子数，则$\{X(t),t\geqslant 0\}$是齐次马尔可夫过程，证明：$X(t)$是纯生过程．

证明　转移概率为

$$\begin{cases} p_{i,i+1}(\Delta t)=i\lambda\Delta t+o(\Delta t), & i\geqslant 0 \\ p_{ii}(\Delta t)=1-i\lambda\Delta t+o(\Delta t), & i\geqslant 0 \\ p_{i,i-1}(\Delta t)=o(\Delta t), & i\geqslant 1 \\ p_{ij}(\Delta t)=o(\Delta t), & |i-j|>1 \end{cases}$$

由此可以得出

$$q_{i,i+1}=i\lambda,\quad i\geqslant 0,\qquad q_{ii}=-i\lambda,\quad i\geqslant 0,$$
$$q_{i,i-1}=0,\quad i\geqslant 1,\qquad q_{ij}=0,\quad |i-j|>1.$$

所以 $X(t)$ 是纯生过程.

【例 5.8】 设 $\{X(t),t\in[0,+\infty)\}$ 是一个有线性生殖率的纯生过程，若 $P\{X(0)=k\}=p(1-p)^{k-1}$，即 $X(0)$ 服从参数为 p 的几何分布($k\geqslant 1$)，求 $E[X(t)]$ 和 $D[X(t)]$.

解 因为 $\{X(t),t\in[0,+\infty)\}$ 是纯生过程,所以有

$$E[X(t+s)-X(s)\mid X(s)=m]=me^{\lambda t}(1-e^{-\lambda t})$$
$$D[X(t+s)-X(s)\mid X(s)=m]=me^{-2\lambda t}(1-e^{-\lambda t})$$
$$D[X(t+s)-X(s)\mid X(s)=m]=E\{[X(t+s)-X(s)]^2\mid X(s)=m\}-\{E[X(t+s)-X(s)\mid X(s)=m]\}^2$$

由此得出

$$\begin{aligned}E[X(t)-X(0)]&=\sum_{n=0}^{+\infty}E[X(t)-X(0)\mid X(0)=n]P\{X(0)=n\}\\&=\sum_{n=1}^{+\infty}ne^{\lambda t}(1-e^{-\lambda t})p(1-p)^{n-1}\\&=e^{\lambda t}p(1-e^{-\lambda t})\sum_{n=1}^{+\infty}n(1-p)^{n-1}\end{aligned}$$

由 $f(x)=\sum_{n=1}^{+\infty}nx^{n-1}=\left(\sum_{n=1}^{+\infty}x^n\right)'=\left(\frac{1}{1-x}\right)'=\frac{1}{(1-x)^2}$,得

$$E[X(t)-X(0)]=p^{-1}e^{\lambda t}(1-e^{-\lambda t})$$

而

$$E[X(0)]=\sum_{n=1}^{+\infty}np(1-p)^{n-1}=p^{-1}$$
$$E[X(t)]=p^{-1}[e^{\lambda t}(1-e^{-\lambda t})+1]=p^{-1}e^{\lambda t}$$
$$\begin{aligned}D[X(t)]&=E[X(t)^2]-\{E[X(t)]\}^2\\&=\sum_{m=1}^{+\infty}\sum_{n=m}^{+\infty}n^2C_{n-1}^{n-m}(e^{\lambda t})^m(1-e^{-\lambda t})^{n-m}-p^{-2}e^{2\lambda t}\end{aligned}$$

【例 5.9】 ($M/M/S$ 排队问题)某货运港有 s 个装卸船台,在时刻 t 到达的货轮数 $N(t)$ 是强度为 λ 的泊松过程. 到达时若有空闲船台,立即进行装卸服务,否则需排队等候直到空出船台后接受服务. 设每条船的服务时间 T 都服从负指数分布,即

$$P(T\leqslant t)=\begin{cases}1-e^{-\mu t}, & t\geqslant 0\\ 0, & t<0\end{cases}$$

并设各船装卸时间相互独立,且与 $N(t)$ 独立. 用 $X(t)$ 表示时刻 t 进港的货船数.

(1) 证明:$\{X(t),t\geqslant 0\}$ 是齐次马尔可夫过程;

(2) 证明：$\{X(t),t\geqslant 0\}$ 是一个生灭过程；

(3) 求其平稳分布；

(4) 对港口的装卸进行定量分析.

证明　(1) 由于在时刻 t 港内的货船数与 $t+\Delta t$ 时到达的货船数是相互独立的(因为泊松分布的独立增量性)，而时刻 t 正在装卸的货船在 $t+\Delta t$ 时是否结束服务也与该船在时刻 t 前已服务的时间无关(因为负指数分布的无记忆性)，所以 $X(t)$ 是一个马尔可夫过程. 又由

$$\begin{aligned}
&P\{X(t+\Delta t)=i\mid X(t)=i\}\\
=&P\{N(\Delta t)=0,A(\Delta t)=0\mid X(t)=i\}\\
&+P\{\bigcup_{k=1}^{\min(i,s)}\{N(\Delta t)=k,A(\Delta t)=k\mid X(t)=i\}\}\\
=&P\{N(\Delta t)=0,\bigcap_{k=1}^{\min(i,s)}\{T_k>\Delta t\mid X(t)=i\}+o(\Delta t)\}\\
=&\begin{cases}e^{-\lambda\Delta t}e^{-i\mu\Delta t}+o(\Delta t)=1-(\lambda+i\mu)\Delta t+o(\Delta t), & i\leqslant s\\ e^{-\lambda\Delta t}e^{-s\mu\Delta t}+o(\Delta t)=1-(\lambda+s\mu)\Delta t+o(\Delta t), & i>s\end{cases}
\end{aligned}$$

类似地，有

$$\begin{aligned}
&P\{X(t+\Delta t)=i-1\mid X(t)=i\}\\
=&\begin{cases}e^{-\lambda\Delta t}(1-e^{-i\mu\Delta t})+o(\Delta t)=i\mu\Delta t+o(\Delta t), & 1\leqslant i\leqslant s\\ e^{-\lambda\Delta t}(1-e^{-s\mu\Delta t})+o(\Delta t)=s\mu\Delta t+o(\Delta t), & i>s\end{cases}
\end{aligned}$$

$$P\{X(t+\Delta t)=i+1\mid X(t)=i\}=\lambda\Delta t+o(\Delta t),i\geqslant 0$$

故由上面三式可知，从 t 到 $t+\Delta t$ 的转移概率与 t 无关，过程是齐次马尔可夫过程.

(2) 由于

$$q_{ii}=\lim_{\Delta t\to 0}\frac{1-p_{ii}(\Delta t)}{\Delta t}=\begin{cases}\lambda+i\mu, & i\leqslant s\\ \lambda+s\mu, & i>s\end{cases}$$

$$q_{i,i+1}=\lambda,\quad i\geqslant 0$$

$$q_{i,i-1}=\begin{cases}i\mu, & i\leqslant s\\ s\mu, & i>s\end{cases}$$

$$q_{ij}=0,\quad |i-j|\geqslant 2,$$

所以，$X(t)$ 是一个生灭过程.

(3) 柯尔莫哥洛夫向后方程为

$$\begin{cases}p'_{ij}=-(\lambda+i\mu)p_{ij}(t)+i\mu p_{i-1,j}(t)+\lambda p_{i+1,j}(t), & 0\leqslant i\leqslant s\\ p'_{ij}=-(\lambda+s\mu)p_{ij}(t)+s\mu p_{i-1,j}(t)+\lambda p_{i+1,j}(t), & i>s\end{cases}$$

因为 $\lambda_i=\lambda,\mu_i=\begin{cases}i\mu, & 0\leqslant i\leqslant s\\ s\mu, & i>s\end{cases}$.

依公式可得

$$\pi_k = \frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k}\pi_0 = \begin{cases} \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k \pi_0, & 0 \leqslant k \leqslant s \\ \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^k \frac{1}{s^{k-s}}\pi_0, & k > s \end{cases}$$

故当$\frac{\lambda}{s\mu} < 1$时，存在平稳分布. $\lambda < s\mu$ 就是单位时间内到达港口的货船平均数少于 s 个船台上装卸完货物的平均货船数. 这时

$$\pi_0 = \left[1 + \sum_{k=1}^{s} \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k + \frac{\lambda}{s!(s\mu-\lambda)}\left(\frac{\lambda}{\mu}\right)^s\right]^{-1}$$

将 π_0 代入 π_k，即得平稳分布 $\boldsymbol{\pi} = (\pi_0, \pi_1, \pi_2, \cdots)$.

(4) 讨论：货船到港无须等待的概率是

$$a_1 = P\{x(t) \leqslant s-1\} = \pi_0 + \pi_1 + \cdots + \pi_{s-1}$$

需等待两条或两条以上货船装卸完的概率是

$$a_2 = P\{x(t) \leqslant s+1\} = 1 - (\pi_0 + \pi_1 + \cdots + \pi_s)$$

在港口内排队等候的货船平均数是

$$b_1 = \sum_{k=1}^{+\infty} k\pi_{s+k}$$

在港口内轮船的平均数是

$$b_2 = \sum_{k=1}^{+\infty} k\pi_k$$

取定几个 s, λ, μ 的数值得一数据表如表 5.1 所示。

表 5.1　数据表

s	λ	μ	a_1	a_2	b_1	b_2
1	18	20	0.1	0.81	8.1	9
1	18	40	0.55	0.2025	0.37	0.818
1	9	20	0.55	0.2025	0.37	0.818
2	18	20	0.72	0.1257	0.23	1.1285
1	18	10	0.147	0.7674	0.7	9.47

从数据表可知，若 λ 相同，装卸效率提高一倍，则货船无须等待的概率可由 0.1 提高到 0.55；而等待的货船平均数从 8.1 条减少到 0.37 条.

5.4　马尔可夫序列与扩散过程

本节仅简单地介绍马尔可夫序列及扩散过程的定义和统计性质，需要进一步

了解的读者请阅读更专业的书籍.

5.4.1　马尔可夫序列

定义 5.7　一个随机变量序列 X_n，若对任意的 n，有

$$F(x_n \mid x_{n-1}, x_{n-2}, \cdots, x_1) = F(x_n \mid x_1)$$

则称 X_n **为马尔可夫序列**，它是时间参数离散、状态空间连续的马尔可夫过程.

马尔可夫序列同样具有“马氏性”，即已知“现在”，则“将来”与“过去”相互独立，

$$f(x_n, x_s \mid x_r) = f(x_n \mid x_r) f(x_s \mid x_r), n > r > s$$

马尔可夫序列的转移概率密度方程为

$$f(x_n \mid x_s) = \int_{-\infty}^{+\infty} f(x_n \mid x_r) f(x_r \mid x_s) \mathrm{d}x_r, n > r > s$$

定义 5.8　如果一个 n 维向量随机序列 $\{x(k), k \in \mathbf{N}^+\}$ 既是正态的，又是马尔可夫序列，则称之为**正态(高斯)-马尔可夫序列**. 其正态性决定了它们的幅度概率分布，而马尔可夫性决定了序列在时间上的传播规律.

设高斯-马尔可夫序列为 $X(n+1) = AX(n) + W(n)$，其中 A 为常数，$W(n)$ 为正态白噪声，其均值为 $m_w(n)$，方差为 $\sigma_w^2(n)$. 若 $X(n)$ 取值为 x_n，则转移概率密度为 $f(x_{n+1} \mid x_n)$，它也是正态的.

正态-马尔可夫序列的数字特征：

条件均值　　$E[X(n+1) \mid x_n] = Ax_n + m_w(n)$

条件方差　　$D[X(n+1) \mid x_n] = \sigma_w^2(n)$

均值　　$E[X(n+1)] = Am_X(n) + m_w(n)$

方差　　$D[X(n+1)] = A^2\sigma_X^2(n) + \sigma_w^2(n)$

协方差

$$\begin{aligned} C_X(n,s) &= E\{[X(n) - m_X(n)][X(s) - m_X(s)]\} \\ &= A^{n-s}\sigma_X^2(s) \end{aligned}$$

如果 $X(n)$ 是平稳序列，则

$$C_X(n,s) = A^{|n-s|}\sigma_X^2$$

【例 5.10】　设 $X_1, X_2, \cdots, X_n, \cdots$ 是独立随机变量，概率密度函数 $f_{x_n}(x) = f_n(x)$，若令

$$Y_1 = X_1, Y_2 = X_1 + X_2, \cdots, Y_n = X_1 + X_2 + \cdots + X_n$$

证明：$Y_1, Y_2, \cdots, Y_n, \cdots$ 是马尔可夫序列.

证明　因为

$$f(y_1, y_2) = f_{Y_2}(y_2 \mid Y_1 = y_1) f_{Y_1}(y_1)$$

又由已知条件得

$$\begin{aligned} f_{Y_2}(y_2 \mid Y_1 = y_1) &= f_{X_1+X_2}(y_2 \mid X_1 = y_1) \\ &= f_{X_2}(X_2 = y_2 - y_1 \mid x_1 = y_1) \\ &= f_{X_2}(y_2 - y_1) \\ &= f_2(y_2 - y_1) \end{aligned}$$

故 $$f(y_1, y_2) = f_1(y_1) f_2(y_2 - y_1)$$

推广到 n 个随机变量有

$$f(y_1, y_2, \cdots, y_n) = f_1(y_1) f_2(y_2 - y_1) \cdots f_n(y_n - y_{n-1})$$

而 $f(y_n \mid y_{n-1}, y_{n-2}, \cdots, y_1) = \dfrac{f(y_1, y_2, \cdots, y_n)}{f(y_1, y_2, \cdots, y_{n-1})} = f_n(y_n - y_{n-1})$

于是知，$f(y_n \mid y_{n-1}, y_{n-2}, \cdots, y_1)$ 与 $y_{n-2}, y_{n-3}, \cdots, y_1$ 无关，所以 $\{Y_n\}$ 是一个马尔可夫序列.

5.4.2 扩散过程

定义 5.9 状态和时间参数均连续变化的马尔可夫过程称为**连续马尔可夫过程**，又称为**扩散过程**. 定义的概率形式表述如下：

设 $\{X(t), t \in T\}$ 是一个随机过程，若对每个 n 和 T 中的 $0 \leqslant t_1 < t_2 < \cdots < t_n$，有

$$F(x_n, t_n \mid x_{n-1}, t_{n-1}; x_{n-2}, t_{n-2}; \cdots; x_1, t_1) = F(x_n, t_n \mid x_{n-1}, t_{n-1})$$

则称 $\{X(t), t \in T\}$ 是扩散过程.

上式也可用概率密度函数等价地表示为

$$f(x_n, t_n \mid x_{n-1}, t_{n-1}; x_{n-2}, t_{n-2}; \cdots; x_1, t_1) = f(x_n, t_n \mid x_{n-1}, t_{n-1})$$

扩散过程的统计特性完全由其一阶、二阶分布函数决定：

$$\begin{aligned} & f(x_1, t_1; x_2, t_2; \cdots; x_n, t_n) \\ &= f_1(x_1, t_1) \prod_{k=1}^{n-1} f_2(x_{k+1}, t_{k+1} \mid x_k, t_k) \\ &= \left[\prod_{k=1}^{n-1} f_2(x_{k+1}, t_{k+1} \mid x_k, t_k)\right] \Big/ \left[\prod_{k=1}^{n-1} f_1(x_1, t_1)\right] \end{aligned}$$

扩散过程的转移概率为

$$P\{s, x; t, A\} = P\{X(t) \in A \mid X(s) = x\}, s, t \in T, s < t$$

其中 A 为正态空间 I 的一个子集. 当 $A = (-\infty, y)$ 时，有

$$P\{s, x; t, (-\infty, y)\} = P\{X(t) < y \mid X(s) = x\}$$

一般地，将 $P\{s, x; t, (-\infty, y)\}$ 记为

$$F(s, x; t, y) = F(t, y \mid s, x) = P\{X(t) < y \mid X(s) = x\}$$

称之为扩散过程的**转移概率分布**.

若 $F(s,x;t,y)$ 关于 y 的导数存在，则

$$f(s,x;t,y)=f(t,y\mid s,x)=\frac{\partial}{\partial y}F(s,x;t,y)$$

称之为扩散过程的**转移概率密度**，则有

$$F(s,x;t,y)=\int_{-\infty}^{y}f(s,x;t,u)\mathrm{d}u$$

若满足 $F(s,x;t,y)=F(t-s;x,y)=F(\tau;x,y)$，则称 $\{X(t),t\in T\}$ 为**齐次的连续马尔可夫过程**.

设 $\{X(t),t\geqslant 0\}$ 的转移概率分布函数是 $F(s,x;t,y)$，由连续性知，当 Δt 很小时，$X(t+\Delta t)-X(t)$ 也很小，即，对任意 $\delta>0$，有

$$\lim_{\Delta t\to 0}P\{|X(t+\Delta t)-X(t)|>\delta\mid X(t)=x\}=0$$

或

$$\lim_{\Delta t\to 0}\int_{|y-x|>\delta}f(t,x;t+\Delta t,y)\mathrm{d}y=0$$

若考虑加强条件，即对任意 $\delta>0,\Delta t>0$，要求

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|y-x|\geqslant\delta}f_y(t,x;t+\Delta t,y)\mathrm{d}y$$

$$=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|y-x|\geqslant\delta}f_y(t-\Delta t,x;t,y)\mathrm{d}y=0 \tag{5.1}$$

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|y-x|<\delta}(y-x)f_y(t,x;t+\Delta t,y)\mathrm{d}y$$

$$=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|y-x|<\delta}(y-x)f_y(t-\Delta t,x;t,y)\mathrm{d}y=a(t,x) \tag{5.2}$$

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|y-x|<\delta}(y-x)^2f_y(t,x;t+\Delta t,y)\mathrm{d}y$$

$$=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|y-x|<\delta}(y-x)^2f_y(t-\Delta t,x;t,y)\mathrm{d}y=b(t,x) \tag{5.3}$$

满足式(5.1) ～ 式(5.3) 的马尔可夫过程是**扩散过程**，其中

$$f(t,x;t+\Delta t,y)\mathrm{d}y=\mathrm{d}F(t,x;t+\Delta t,y)$$

称 $a(t,x)$ 为**偏移系数**，称 $b(t,x)$ 为**扩散系数**，均不依赖于 δ.

扩散过程的 C-K 方程为

$$f(x_n,t_n\mid x_s,t_s)=\int_{-\infty}^{+\infty}f(x_n,t_n\mid x_r,t_r)f(x_r,t_r\mid x_s,t_s)\mathrm{d}x$$

其中 $t_n>t_r>t_s$.

【例 5.11】　已知一个独立随机变量序列 $X_1,X_2,\cdots,X_n,\cdots$，有

$$f_{X_n}(x)=f_n(x)$$

令 $Y_1=X_1,Y_2=X_1+X_2,\cdots,Y_n=X_1+X_2+\cdots+X_n$，若 $E(X_n)=0$，X_n 与 Y_{n-1} 相互独立，证明：

$$E(Y_n \mid Y_{n-1},Y_{n-2},\cdots,Y_1)=Y_{n-1}$$

证明　由例 5.10 知，$\{Y_n\}$是一个马尔可夫序列，则

$$\begin{aligned}&E(Y_n \mid Y_{n-1},Y_{n-2},\cdots,Y_1)\\&=\int_{-\infty}^{+\infty} y_n f(y_n \mid y_{n-1},y_{n-2},\cdots,y_1)\mathrm{d}y_n\\&=\int_{-\infty}^{+\infty} y_n f(y_n \mid y_{n-1})\mathrm{d}y_n=E(Y_n \mid Y_{n-1})\end{aligned}$$

又因为

$$Y_n=\sum_{i=1}^{n}X_i=X_n+\sum_{i=1}^{n-1}X_i=X_n+Y_{n-1}$$

故

$$\begin{aligned}&E(Y_n \mid Y_{n-1})=E(X_n+Y_{n-1} \mid Y_{n-1})\\&=E(X_n \mid Y_{n-1})+E(Y_{n-1} \mid Y_{n-1})=E(Y_{n-1} \mid Y_{n-1})\end{aligned}$$

设 X 是任一随机变量，$g(x)$ 是一已知函数，由条件数学期望的定义，有

$$E[g(X) \mid x]=\int_{-\infty}^{+\infty} g(x)p[g(x) \mid x]\mathrm{d}x$$

所以

$$P\{g(x) \mid X=x\}=1,P\{g(x) \mid X\neq x\}=0$$

故

$$E[g(x) \mid x]=\int_{-\infty}^{+\infty} g(x)\delta(X-x)\mathrm{d}x=g(X)$$

若令 $X=Y_{n-1}$，$g(X)=Y_{n-1}$. 则有 $E[Y_{n-1} \mid Y_{n-1}]=Y_{n-1}$，从而证得

$$E[Y_n \mid Y_{n-1},Y_{n-2},\cdots,Y_1]=Y_{n-1}$$

【例 5.12】　设$\{X(t),t\in[0,+\infty)\}$ 的转移函数是

$$P\{s,x;t,y\}\triangleq\begin{cases}\dfrac{1}{\sigma\sqrt{2\pi(t-s)}}\displaystyle\int_{-\infty}^{y}\exp\left\{-\dfrac{(y-x)^2}{2\sigma^2(t-s)}\right\}^2\mathrm{d}y,0\leqslant s<t\\ I_{(-\infty|y)}(x), & s=t\end{cases}$$

证明：此马尔可夫过程是扩散过程，并求出偏移系数 $a(t,x)$ 和扩散系数 $b(t,x)$.

证明　因为

$$\begin{aligned}&\frac{1}{\Delta t}\int_{-\infty}^{+\infty}(x-y)^4 f(t,x;t+\Delta t,y)\mathrm{d}y\\&=\frac{1}{\Delta t}\int_{-\infty}^{+\infty}(x-y)^4\frac{1}{\sigma\sqrt{2\pi(t-s)}}\int_{-\infty}^{y}\exp\left\{-\frac{(y-x)^2}{2\sigma^2(t-s)}\right\}^2\mathrm{d}y=3\sigma^4\Delta t\end{aligned}$$

于是对于固定的 δ 及 $i=0,1,2$，有

$$\begin{aligned}&\int_{|y-x|>\delta}(x-y)^i f(t,x;t+\Delta t,y)\mathrm{d}y\\&\leqslant\frac{1}{\delta^{4-i}}\int_{-\infty}^{+\infty}(x-y)^4 f(t,x;t+\Delta x,y)\mathrm{d}y=\frac{1}{\delta^{4-i}}3\sigma^4(\Delta t)^2\end{aligned}$$

$$a(t,x)=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|y-x|\leqslant\delta}(y-x)f(t,x;t+\Delta t,y)\mathrm{d}y$$
$$=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{-\infty}^{+\infty}(y-x)f(t,x;t+\Delta t,y)\mathrm{d}y=0$$
$$b(t,x)=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|y-x|\leqslant\delta}(y-x)^2f(t,x;t+\Delta t,y)\mathrm{d}y$$
$$=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{-\infty}^{+\infty}(y-x)^2f(t,x;t+\Delta t,y)\mathrm{d}y=\delta^2$$

所以，扩散过程的三个条件都满足，此马尔可夫过程是扩散过程，且偏移系数 $a(t,x)=0$，扩散系数 $b(t,x)=\delta^2$.

5.5　应用举例

在实际应用问题的建模中，连续时间马尔可夫链使用非常广泛，下面举一些在应用中具有一定代表性的例题，可窥一斑.

【例 5.13】（Yule 过程）考察生物群体繁殖过程模型. 设群体中各个生物体的繁殖都是相互独立的、强度为 λ 的 Poisson 过程，并且群体中没有死亡，此过程称为 Yule 过程.

(1) 说明 Yule 过程是连续时间的 Markov 链；

(2) 计算 Yule 过程的转移概率.

解　(1) 设时刻 0 群体中有 1 个个体，则群体将有的个体数为 $\{1,2,3,\cdots\}$. 以 $T_i(i\geqslant 1)$ 记群体数目从 i 增加到 $i+1$ 所需的时间，由 Yule 过程定义知，当群体数目为 i 时，这 i 个个体是以相互独立的 Poisson 过程来产生后代的，由 Poisson 过程的可加性（习题 3.3）知，这相当于强度为 λi 的 Poisson. 由 Poisson 过程的独立增量性易知，T_i 与状态的转移是独立的，且 T_i 是相互独立的、参数为 λi 的指数随机变量，这说明 Yule 过程是连续时间的 Markov 过程.

(2) 下面求 Yule 过程的转移概率：

由于考虑的是单一的始祖 $X(0)=1$，根据模型的假定有

$$p_{kk}(h)=P\{X(t+h)-X(t)=0\mid X(t)=k\}=1-k\lambda h+o(h)$$
$$p_{k,k+1}(h)=P\{X(t+h)-X(t)=1\mid X(t)=k\}=k\lambda h+o(h)$$
$$p_{k,k+i}(h)=P\{X(t+h)-X(t)=i\mid X(t)=k\}=o(h),i\geqslant 2$$
$$P\{X(t+h)-X(t)<0\mid X(t)=k\}=0$$

由此导出 $p_{ij}(t)$ 满足的微分方程，首先

$$\lim_{h\to 0}\frac{1-p_{kk}(h)}{h}=q_{kk}=k\lambda$$

$$\lim_{h\to 0}\frac{p_{k,k+1}(h)}{h}=q_{k,k+1}=k\lambda$$

从而

$$\begin{aligned}p'_{ij}(t)&=q_{i,i+1}p_{i+1,j}(t)-q_{ii}p_{ij}(t)\\&=\lambda i p_{i+1,j}(t)-\lambda i p_{ij}(t)\end{aligned}$$

当 $j=i$ 时，

$$p'_{ii}=-\lambda i p_{ii}(t)$$

解得

$$p_{ii}(t)=\mathrm{e}^{-\lambda it}$$

当 $j=i+1$ 时，

$$\begin{aligned}p'_{i,i+1}(t)&=\lambda(i+1)p_{i+1,i+1}(t)-i\lambda p_{i,i+1}(t)\\&=\lambda(i+1)\mathrm{e}^{-\lambda(i+1)t}-i\lambda p_{i,i+1}(t)\end{aligned}$$

这是一个一阶线性微分方程，它的解为

$$p_{i,i+1}(t)=i\mathrm{e}^{-\lambda it}(1-\mathrm{e}^{-\lambda t})$$

当 $j=i+2$ 时，

$$\begin{aligned}p'_{i,i+2}(t)&=\lambda i p_{i+1,i+2}(t)-i\lambda p_{i,i+2}(t)\\&=[\lambda(i+1)]^2 t\mathrm{e}^{-\lambda(i+1)t}-i\lambda p_{i,i+2}(t)\end{aligned}$$

解得

$$p_{i,i+2}(t)=i(i-1)\mathrm{e}^{-\lambda it}(1-\mathrm{e}^{-\lambda t})^2$$

依次类推，

$$p_{ij}(t)=\binom{j-1}{i-1}\mathrm{e}^{-\lambda ti}(1-\mathrm{e}^{-\lambda t})^{i},\quad j\geqslant i\geqslant 1$$

【例 5.14】（排对问题）设有一服务台，$[0,t)$ 内到达服务台的顾客数是服从 Poisson 分布的随机变量. 单位时间到达服务台的平均人数是 λ. 服务台只有一个服务员，对顾客的服务时间是按负指数分布的随机变量，平均服务时间为 $1/\mu$. 如果服务台空闲，则到达的顾客立刻得到服务；如果顾客到达时服务员正在为另一顾客服务，则他必须排队等候；如果顾客到达时发现已经有两人在等候，则他就离开不再回来. 设 $X(t)$ 代表在 t 时刻系统内顾客人数（包括正在被服务的顾客和排队等候的顾客）. 假设系统在 $t=0$ 时处于零状态，即服务人员空闲. 求 t 时刻系统处于状态 j 的无条件概率 $p_j(t)$ 所满足的微分方程.

解　(1) 写出状态空间：$I=\{0,1,2,3\}$.

(2) 求 Q 矩阵：

① 当 $X(t)=0$ 时，在 $[t,t+\Delta t)$ 内到达一个顾客的概率为

$$p_{01}(\Delta t)=\lambda\Delta t+o(\Delta t)$$

在 $[t,t+\Delta t)$ 内到达两个或两个以上顾客的概率为

$$p_{0j}(\Delta t)=o(\Delta t)\quad j=2,3$$

因此

$$q_{01}=\lim_{\Delta t\to 0}\frac{p_{01}(\Delta t)}{\Delta t}=\lambda$$

$$q_{0j}=\lim_{\Delta t\to 0}\frac{p_{0j}(\Delta t)}{\Delta t}=0$$

由$\sum_{j\in I}q_{ij}=0$，可得$q_{00}=-\lambda$.

② 当$X(t)=1$时，表示在t时刻有一个顾客正在被服务. 由指数分布的无记忆性可知，在$[t,t+\Delta t)$内完成服务的概率为

$$(1-e^{-\mu\Delta t})=\mu\Delta t+o(\Delta t)$$

由此可知，在Δt时间内系统由1状态转入到0状态的概率为

$$p_{10}(\Delta t)=[\mu\Delta t+o(\Delta t)][1-\lambda\Delta t+o(\Delta t)]=\mu\Delta t+o(\Delta t)$$

故

$$q_{10}=\lim_{\Delta t\to 0}\frac{p_{10}(\Delta t)}{\Delta t}=\mu$$

在Δt时间内系统由1状态转入到2状态的概率为

$$p_{12}(\Delta t)=[1-\mu\Delta t+o(\Delta t)][\lambda\Delta t+o(\Delta t)]=\lambda\Delta t+o(\Delta t)$$

故

$$q_{12}=\lim_{\Delta t\to 0}\frac{p_{12}(\Delta t)}{\Delta t}=\lambda$$

同理可得

$$q_{13}=0$$
$$q_{11}=-(\lambda+\mu)$$
$$q_{20}=0$$
$$q_{21}=\mu$$
$$q_{23}=\lambda$$
$$q_{22}=-(\lambda+\mu)$$

③ 当$X(t)=3$时，系统不再接受新顾客，此时，状态只可转到2或仍在3. 当$X(t)=3$时，在Δt时间内完成服务的概率为

$$(1-e^{-\mu\Delta t})=\mu\Delta t+o(\Delta t)$$

因此

$$q_{32}=\mu$$
$$q_{30}=0$$
$$q_{31}=0$$
$$q_{33}=-\mu$$

于是可得Q矩阵如下：

$$Q=\begin{bmatrix}-\lambda & \lambda & 0 & 0\\ \mu & -(\lambda+\mu) & \lambda & 0\\ 0 & \mu & -(\lambda+\mu) & \lambda\\ 0 & 0 & \mu & -\mu\end{bmatrix}$$

(3) 写出方程

$$\begin{cases}\dfrac{\mathrm{d}p_0(t)}{\mathrm{d}t}=-\lambda p_0(t)+\mu p_1(t)\\ \dfrac{\mathrm{d}p_1(t)}{\mathrm{d}t}=\lambda p_0(t)-(\lambda+\mu)p_1(t)+\mu p_2(t)\\ \dfrac{\mathrm{d}p_2(t)}{\mathrm{d}t}=\lambda p_1(t)-(\lambda+\mu)p_2(t)+\mu p_3(t)\\ \dfrac{\mathrm{d}p_3(t)}{\mathrm{d}t}=\lambda p_2(t)-\mu p_3(t)\end{cases}$$

初始条件为

$$\begin{cases}p_0(0)=1\\ p_j(0)=0,j=1,2,3\end{cases}$$

【例 5.15】 (生灭过程)仍然考虑一个生物群体的繁殖模型.每个个体生育后代如 Yule 过程的假定,但是每个个体将以指数速率死亡,这就是一个生灭过程.它的状态为{0, 1, 2,…},在状态 i,它能够转移到 $i+1$ 或 $i-1$. 设 T_i 表示过程从 i 到达 $i+1$ 或 $i-1$ 的时间,类似于 Yule 过程,T_i 服从参数为 $i\mu+i\lambda$ 的指数分布,并且下次转移到 $i+1$ 的概率是$\dfrac{\lambda}{\mu+\lambda}$,转移到 $i-1$ 的概率为$\dfrac{\mu}{\lambda+\mu}$. 这说明生灭过程是连续时间的 Markov 过程.

同样的,生灭过程也满足

$$p_{k,k+1}(h)=P\{X(t+h)-X(t)=1\mid X(t)=k\}=k\lambda h+o(h)$$

$$p_{k,k-1}(h)=P\{X(t+h)-X(t)=-1\mid X(t)=k\}=k\mu h+o(h)$$

$$p_{kk}(h)=P\{X(t+h)-X(t)=i\mid X(t)=k\}=1-(\lambda+\mu)kh+o(h),i\geqslant 2$$

$$p_{ij}(0)=\delta_{ij}$$

从而导出生灭过程的 Q 矩阵为

$$Q=\begin{bmatrix}-\lambda & \lambda & \cdots & \cdots\\ \mu & -(\lambda+\mu) & \lambda & \cdots\\ \cdots & 2\mu & -(\lambda+2\mu) & \lambda\\ \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots\end{bmatrix}$$

向后方程为

$$p'_{ij}(t)=i\mu p_{i-1,j}(t)-(\lambda+i\mu)p_{ij}(t)+\lambda p_{i+1,j}(t),i\geqslant 1$$

$$p'_{0j}(t) = -\lambda p_{0j}(t) + \lambda p_{1j}(t)$$

向前方程为

$$p'_{ij}(t) = (i+1)\mu p_{i,j+1} - (\lambda + i\mu)p_{ij}(t) + \lambda p_{i,j-1}(t)$$
$$p'_{i0}(t) = -\lambda p_{i0}(t) + \mu p_{i1}(t)$$

【例 5.16】 (系统调整问题)考察某一系统运作情况,如果系统运作正常,则认为系统处于状态1;如果系统正在调整,则认为系统处于状态0.系统运作一段时间后,会遇到不正常运作,此时需要调整,调整后恢复运作.假定系统从开始运作直至需要调整的运作时间是随机的,服从参数为μ的指数分布,密度函数为$\mu e^{-\mu t}$,$t>0$;而调整期也是随机的,服从参数为λ的指数分布,密度函数为$\lambda e^{-\lambda t}$,$t>0$.假定运作周期是相互独立的,调整期也是相互独立的.如果令$X(t)$为系统在时刻t所处的状态,则$X(t)$是时齐Markov链.试计算$X(t)$的转移概率.

解 下面列出Kolmogorov微分方程.如果系统在时刻t处于状态0,而在时刻$t+\Delta t$变为状态1,当时间增量Δt很小时,

$$p_{01}(\Delta t) = \int_0^{\Delta t} \lambda e^{-\lambda t} dt = 1 - e^{-\lambda \Delta t} = \lambda \Delta t + o(\Delta t)$$

所以

$$q_{01} = \lim_{\Delta t \to 0} \frac{p_{01}(\Delta t)}{\Delta t} = \lambda$$

同理

$$p_{10}(\Delta t) = \mu \Delta t + o(\Delta t), \text{则 } q_{10} = \mu$$

于是得到

$$\boldsymbol{Q} = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$$

从而Kolmogorov向前方程为

$$p'_{i0}(t) = -\lambda p_{i0}(t) + \mu p_{i1}(t), \quad i = 0,1$$
$$p'_{i1}(t) = \lambda p_{i0}(t) - \mu p_{i1}(t), \quad i = 0,1$$

由于$p_{i0}(t) + p_{i1}(t) = 1$,故将$p_{i1}(t) = 1 - p_{i0}(t)$代入第一个方程得

$$p'_{i0}(t) + (\lambda + \mu)p_{i0}(t) = \mu$$

在初始条件$p_{ij}(0) = \delta_{ij}$下,上述方程的解为

$$y = e^{-\int p(x)dx}\left[c + \int q(x) e^{\int p(x)dx} dx\right]$$

其中$p = \lambda + \mu, q = \mu$,解得

$$p_{i0}(t) = \frac{\mu}{\lambda + \mu} + c e^{-(\lambda+\mu)t}$$

利用$p_{00}(0) = 1, p_{10}(0) = 0$确定常数$c$,得

$$p_{00}(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t}$$

$$p_{10}(t)=\frac{\lambda}{\lambda+\mu}-\frac{\mu}{\lambda+\mu}\mathrm{e}^{-(\lambda+\mu)t}$$

再利用 $p_{i1}(t)=1-p_{i0}(t)$ 得

$$p_{01}(t)=\frac{\lambda}{\lambda+\mu}-\frac{\lambda}{\lambda+\mu}\mathrm{e}^{-(\lambda+\mu)t}$$

$$p_{11}(t)=\frac{\mu}{\lambda+\mu}+\frac{\mu}{\lambda+\mu}\mathrm{e}^{-(\lambda+\mu)t}$$

【例 5.17】（传染模型）考虑有 m 个个体的群体，在时刻 0 由一个已被感染的个体与 $m-1$ 个未受到感染但可能被感染的个体组成. 个体一旦受感染将永远处于此状态. 假设在任意长为 h 的时间区间内，任意一个已被感染的个体将以概率 $ah+o(h)$ 引起任一指定的未感染者成为感染者. 若以 $X(t)$ 记时刻 t 群体中已受感染的个体数，则 $\{X(t),t\geqslant 0\}$ 是一纯生过程，其中

$$\lambda_n=\begin{cases}(m-n)na, & n=1,2,\cdots,m-1\\ 0, & \text{其他}\end{cases}$$

这是因为当有 n 个已被感染的个体时，则 $m-n$ 个未受到感染的每一个个体将以速率 na 变成已感染者.

记 T 为直至整个群体被感染的时间，T_i 为从 i 个已感染者到 $i+1$ 个已感染者的时间，则

$$T=\sum_{i=1}^{m-1}T_i$$

由于 T_i 是相互独立的指数随机变量，其参数分别为

$$\lambda_i=(m-i)ia,\quad i=1,2,\cdots,m-1$$

所以，

$$E(T)=\sum_{i=1}^{m-1}E(T_i)=\frac{1}{a}\sum_{i=1}^{m-1}\frac{1}{i(m-i)}$$

$$D(T)=\sum_{i=1}^{m-1}D(T_i)=\frac{1}{a^2}\sum_{i=1}^{m-1}\left[\frac{1}{i(m-i)}\right]^2$$

习　题　5

1. 以 Yule 过程计算群体总数从 1 增长到 N 的平均时间.
2. 假定生物群体中的各个个体以指数率 λ 出生，以指数率 μ 死亡，另外还存在由迁入引起的指数增长率 θ，试对此建立一个生灭模型.
3. 设有一个质点在 1,2,3 上做随机跳跃，在时刻 t 它位于三点之一，且在 $[t,t+h]$ 内以概率 $0.5h+o(h)$ 分别可以跳到其他两个状态. 试求转移概率满足的 Kolmogorov-Feller 方程，转移概率 $p_{ij}(t)$ 及平稳分布.

4. 设$[0,t]$内到达的顾客数服从泊松分布，参数为λt. 设有单个服务员，服务时间为指数分布的排队系统($M/M/1$)，平均服务时间为$1/\mu$. 试证明：

   (1) 在服务员的服务时间内到达的顾客的平均数为λ/μ；

   (2) 在服务员的服务时间内无顾客到达的概率为$\mu/(\lambda+\mu)$.

5. (机器维修问题) 设某机器的正常工作时间是一负指数分布的随机变量，它的平均正常工作时间为$1/\lambda$；它损坏后的修复时间也是一负指数分布的随机变量，它的平均修复时间为$1/\mu$. 如果该机器在$t=0$时是正常工作的，问在$t=10$时该机器处于正常工作状态的概率是多少？长时间工作下去，机器处于正常状态的概率如何？

6. (电话交换问题) 某电话总机有n条线路. 在某一呼唤来到时如有空闲线路，则该呼唤占用其中某一条空闲线路，并开始通话. 如果通话结束，则该线路使用完毕而称为空闲线路，等待下一次呼唤. 如果呼唤来到时遇到n条线路均被占用，则该呼唤遭到拒绝而消失. 设有按Poisson分布的呼唤流，即在$[t,t+\Delta t)$内来到一次呼唤的概率为$\lambda\Delta t+o(\Delta t)$，来到两次或两次以上呼唤的概率为$o(\Delta t)$；并设如果某一线路在某时刻$t$被占用，而在$[t,t+\Delta t)$内这条线路空闲出来的概率为$\mu\Delta t+o(\Delta t)$，即通话时间按负指数分布. 求总机在$t$时刻有$k$条线路被占用的概率，以及当$t\to+\infty$时，有$k$条线路被占用的概率.

第6章 鞅和布朗运动

6.1 鞅的基本概念和性质

鞅是解决许多概率理论及其应用问题的重要工具. 比如随机游动、随机点过程、精算风险分析和数理金融等. 比较典型的随机模型有“公平赌博”,即赌博中,每个“参与者”有相同的机会赢得和失去. 鞅作为一类特殊的随机过程由 J. Ville 和 P. Levy 引入,1953 年 J. L. Doob 认识到鞅具有很重要的理论和实用价值,类似于马尔可夫过程,进而基于离散和连续参数空间 T 给出了离散时间鞅和连续时间鞅的定义.

【例 6.1】 公平赌博.

设$\{Y_n, n \geqslant 1\}$i. i. d,表示输和赢,$P\{Y_n = 1\} = P\{Y_n = -1\} = 0.5$,赌注为 $b_n = b_n(Y_1, Y_2, \cdots, Y_{n-1})$,设 X_0 表示赌博前的初始赌资,则第 n 次赌博后的赌资 X_n 为

$$X_n = X_0 + \sum_{i=1}^{n} b_i Y_i$$

则

$$E(X_{n+1} \mid Y_1, Y_2, \cdots, Y_n) = X_n$$

即已知前面 n 次赌博的结果,第 $n+1$ 次赌博后的平均赌资等于第 n 次赌博后的赌资.

证明 由于 $X_{n+1} = X_n + b_{n+1} Y_{n+1}$

$$\begin{aligned} E(X_{n+1} \mid Y_1, Y_2, \cdots, Y_n) &= E(X_n + b_{n+1} Y_{n+1} \mid Y_1, Y_2, \cdots, Y_n) \\ &= X_n + E(b_{n+1} Y_{n+1} \mid Y_1, Y_2, \cdots, Y_n) \\ &= X_n + b_{n+1} E(Y_{n+1} \mid Y_1, Y_2, \cdots, Y_n) \end{aligned}$$

由于$\{Y_n, n \geqslant 1\}$i. i. d,$E(Y_n) = 0, n = 1, 2, \cdots$,因此

$$E(Y_{n+1} \mid Y_1, Y_2, \cdots, Y_n) = E(Y_{n+1}) = 0$$

从而

$$E(X_{n+1} \mid Y_1, Y_2, \cdots, Y_n) = X_n$$

定义 6.1 随机过程$\{X_n, n \geq 0\}$称为关于$\{Y_n, n \geq 0\}$的**鞅**，如果满足

(1)X_n 是$(Y_0, Y_1, \cdots, Y_n)$的函数；

(2)$E(|X_n|) < +\infty$；

(3)$E(X_{n+1} \mid Y_0, Y_1, \cdots, Y_n) = X_n$.

定义 6.2 称$\{X_n, n \geq 0\}$是关于$\{Y_n, n \geq 0\}$的**上鞅**，如果对任意$n \geq 0$，X_n是$(Y_0, Y_1, \cdots, Y_n)$的函数，$E(X_n^-) < +\infty$，并且$E(X_{n+1} \mid Y_0, Y_1, \cdots, Y_n) \leq X_n$. 类似地，称$\{X_n, n \geq 0\}$是关于$\{Y_n, n \geq 0\}$的**下鞅**，如果对任意$n \geq 0$，$X_n$是$(Y_0, Y_1, \cdots, Y_n)$的函数，$E(X_n^+) < +\infty$，并且$E(X_{n+1} \mid Y_0, Y_1, \cdots, Y_n) \geq X_n$.

这里$X_n^+ = \max\{0, X_n\}$，$X_n^- = \max\{0, -X_n\}$.

显然，鞅的定义描述的是"公平赌博"，上、下鞅分别描述了"不利赌博"和"有利赌博".

下面定义σ代数的鞅，令$\mathscr{F}_n = \sigma(Y_0, Y_1, \cdots, Y_n)$，$X_n$关于$\mathscr{F}_n$可测，则

$$E(X_{n+1} \mid Y_0, Y_1, \cdots, Y_n) = X_n \Leftrightarrow E(X_{n+1} \mid \mathscr{F}_n) = X_n$$

定义 6.3 设$(\Omega, \mathscr{F}, P)$是概率空间，称$\{\mathscr{F}_n, n \geq 1\}$是$\mathscr{F}$的**非降子$\sigma$代数流**，若$\{\mathscr{F}_n, n \geq 1\}$是$\mathscr{F}$的子$\sigma$代数，且$\mathscr{F}_n \subset \mathscr{F}_{n+1}$；称随机序列$\{X_n, n \geq 0\}$是$\{\mathscr{F}_n, n \geq 1\}$**适应的**，若$X_n$关于$\mathscr{F}_n$可测.

定义 6.4 设$\{\mathscr{F}_n, n \geq 1\}$是$\mathscr{F}$的非降子$\sigma$代数流，随机序列$\{X_n, n \geq 0\}$称为是关于$\{\mathscr{F}_n, n \geq 1\}$的**鞅**[简记为$(X_n, \mathscr{F}_n)$是鞅]，如果

(1)$\{X_n, n \geq 0\}$是$\{\mathscr{F}_n, n \geq 1\}$适应的；

(2)任意n，$E(|X_n|) < +\infty$；

(3)$E(X_{n+1} \mid \mathscr{F}_n) = X_n$, a.s.

定义 6.5 设$\{\mathscr{F}_n, n \geq 1\}$是$\mathscr{F}$的非降子$\sigma$代数流，随机序列$\{X_n, n \geq 0\}$称为是关于$\{\mathscr{F}_n, n \geq 1\}$的**下(上)鞅**[简记为$(X_n, \mathscr{F}_n)$是下(上)鞅]，如果

(1)$\{X_n, n \geq 0\}$是$\{\mathscr{F}_n, n \geq 1\}$适应的；

(2)任意n，$E(|X_n|) < +\infty$；

(3)$E(X_{n+1} \mid \mathscr{F}_n) \geq (\leq) X_n$, a.s.

如果考虑连续参数空间T，对于连续时间随机过程$\{X_t, t \geq 0\}$，有如下的连续时间鞅的定义：

定义 6.6 设$\{\mathscr{F}_t, t \geq 0\}$是$\mathscr{F}$的非降子$\sigma$代数流(即$\mathscr{F}_s \subset \mathscr{F}_t, s < t$)，随机过程$\{X_t, t \geq 0\}$称为是关于$\{\mathscr{F}_t, t \geq 0\}$的**鞅**[简记为$(X_t, \mathscr{F}_t)$是鞅]，如果

(1)$\{X_t, t \geq 0\}$是$\{\mathscr{F}_t, t \geq 0\}$适应的；

(2)任意t，$E(|X_t|) < +\infty$；

(3)对一切$0 \leq s < t$，$E(X_t \mid \mathscr{F}_s) = X_s$, a.s.

定义 6.7 设$\{\mathscr{F}_t,t\geqslant 0\}$是$\mathscr{F}$的非降子$\sigma$代数流，随机过程$\{X_t,t\geqslant 0\}$称为是关于$\{\mathscr{F}_t,t\geqslant 0\}$的下(上)鞅[简记为$(X_t,\mathscr{F}_t)$是下(上)鞅]，如果

(1)$\{X_t,t\geqslant 0\}$是$\{\mathscr{F}_t,t\geqslant 0\}$适应的；

(2) 任意t，$E(|X_t|)<+\infty$；

(3)$E(X_t\mid\mathscr{F}_s)\geqslant(\leqslant)X_s$，a.s.

不加证明地给出如下连续时间鞅的基本性质：

(1) 若$(X_t,\mathscr{F}_t)$是鞅，则$E(X_t)=E(X_0)$，$\forall t\in T$.

(2) 若$(X_t,\mathscr{F}_t)$是下鞅$\Leftrightarrow(-X_t,\mathscr{F}_t)$是上鞅.

(3) 若$(X_t,\mathscr{F}_t)$，$(Y_t,\mathscr{F}_t)$是下(上)鞅，则对任意$a>0,b>0$，$(aX_t+bY_t,\mathscr{F}_t)$是下(上)鞅.

(4) 若$(X_t,\mathscr{F}_t)$，$(Y_t,\mathscr{F}_t)$是下鞅，则$(\max\{X_t,Y_t\},\mathscr{F}_t)$是下鞅.

(5) 设$(X_t,\mathscr{F}_t)$是鞅(下鞅)，$\varphi(x)$是$\mathbf{R}$上的凸函数，且满足$E[\varphi(X_t)^+]<+\infty$，则$\varphi(X_t)$是关于$\{\mathscr{F}_t,t\geqslant 0\}$的下鞅.

如果考虑离散参数空间T，则离散鞅具有类似上述性质(1)～(5)，不再另外给出.

【例 6.2】 设$(\Omega,\mathscr{F},P)$为一概率空间，v为$\mathscr{F}$上的一有限测度，$\mathscr{F}_1,\mathscr{F}_2,\cdots$为$\mathscr{F}$上的非降子$\sigma$代数流. 假定$v$和$P$都限制在$\mathscr{F}_n$上时，$v$受控于$P$，即假定若$A\in\mathscr{F}_n$且$P\{A\}=0$则有$v(A)=0$. 由 Radon-Nikodym 定理得知，当$v,P$都限制在$\mathscr{F}_n$时，存在$v$对于$P$的 Radon-Nikodym 导数，记为$X_n$. 则$X_n$关于$\mathscr{F}_n$可测且对$P$可积，对任意$A\in\mathscr{F}_n$，$X_n$满足

$$v(A)=\int_A X_n\mathrm{d}P$$

又由于$A\in\mathscr{F}_n$，因此$A\in\mathscr{F}_{n+1}$，故$v(A)=\int_A X_{n+1}\mathrm{d}P$. 所以，由条件期望的定义知

$$E(X_{n+1}\mid\mathscr{F}_n)=X_n$$

【例 6.3】 设Y是$(\Omega,\mathscr{F},P)$上的可积随机变量，$\{\mathscr{F}_n\}$为$\mathscr{F}$的非降子σ代数流，定义

$$X_n=E(Y\mid\mathscr{F}_n)$$

即X_n为Y关于$\mathscr{F}_n$的条件期望. 由条件期望的定义知，X_n关于$\mathscr{F}_n$可测且可积. 且由条件期望的性质知

$$E(X_{n+1}\mid\mathscr{F}_n)=E[E(Y\mid\mathscr{F}_{n+1})\mid\mathscr{F}_n]=E(Y\mid\mathscr{F}_n)=X_n$$

因此，$\{X_n,\mathscr{F}_n\}$为鞅.

【例 6.4】 设X_n，i.i.d，$E(|X_n|)<+\infty$，证明：$S_0=0$，$S_n=\sum_{k=1}^{n}X_k$是关于$(X_1,X_2,\cdots,X_n)$的鞅.

证明 显然S_n关于$(X_1,X_2,\cdots,X_n)$可测，且

$$E(S_n)=E\left(\sum_{k=1}^{n}X_k\right)=\sum_{k=1}^{n}E(X_k)<+\infty$$

令 $\mathscr{F}_n=\sigma(X_1,X_2,\cdots,X_n)$，则

$$\begin{aligned}E(S_{n+1}\mid\mathscr{F}_n)&=E(X_{n+1}+S_n\mid\mathscr{F}_n)\\&=E(X_{n+1}\mid\mathscr{F}_n)+E(S_n\mid\mathscr{F}_n)\\&=E(X_{n+1})+S_n=S_n\end{aligned}$$

综上所述，$S_n=\sum\limits_{k=1}^{n}X_k$ 是关于$(X_1,X_2,\cdots,X_n)$的鞅.

注意：若 $\mu=E(X_n)\geqslant(\leqslant)0$，则

(1) $\{S_n,n\geqslant 0\}$ 是关于$(X_1,X_2,\cdots,X_n)$的下(上)鞅.

(2) $\{M_n=S_n-n\mu\}$ 是关于$(X_1,X_2,\cdots,X_n)$的鞅.

【例 6.5】 考虑一个公平博弈的问题，设 $X_1,X_2,\cdots,X_n,\cdots$ 表示每次赌博(投硬币)的结果，则它们独立同分布，且

$$P\{X_i=1\}=P\{X_i=-1\}=\frac{1}{2}$$

假设按照以下规则来赌博，每次投掷硬币之前的赌注都比上一次翻一倍，直到赢了赌博即停. 令 W_n 表示第 n 次赌博后的总钱数，$W_0=0$.

假设前 n 次投出的硬币都出现了反面，按照规则，已经输了 $1+2+4+\cdots+2^{n-1}=(2^n-1)$ 元，即 $W_n=-(2^n-1)$，假如下一次硬币出现正面，按规则 $W_{n+1}=2^n-(2^n-1)=1$，由公平的前提知

$$P\{W_{n+1}=1\mid W_n=-(2^n-1)\}=\frac{1}{2}$$

$$P\{W_{n+1}=-2^n-2^n+1\mid W_n=-(2^n-1)\}=\frac{1}{2}$$

无论何时，只要赢了就停止赌博，从而 W_n 从赢了以后就不再变化，即

$$P\{W_{n+1}=1\mid W_n=1\}=1$$

易证 $E\{W_{n+1}\mid\mathscr{F}_n\}=W_n$，即 $\{W_n,\mathscr{F}_n\}$ 是鞅.

【例 6.6】 (Polyas 坛子抽样模型) 考虑一个装有红黄两色球的坛子. 假设最初坛子中装有红、黄色球各一个，每次都按如下规则有放回地随机抽取：如果拿出的是红色的球，则放回的同时再加入一个同色的球，如果拿出的是黄色的球也按同样方式操作. 设 M_n 表示第 n 次抽取后的坛子中的红球所占的比例，证明：$\{M_n,n\geqslant 1\}$ 是鞅.

证明 第 n 次抽取后坛子里球的个数是 $n+2$，以 X_n 表示第 n 次抽取后的坛子中的红球个数，M_n 表示第 n 次抽取后的坛子中的红球所占的比例，则 $M_n=\dfrac{X_n}{n+2}$. $\{X_n,n\geqslant 1\}$ 是一个非时齐的 Markov 链，其转移概率为

$$P\{X_{n+1}=k+1 \mid X_n=k\}=\frac{k}{n+2}$$

$$P\{X_{n+1}=k \mid X_n=k\}=\frac{n+2-k}{n+2}$$

因此

$$\begin{aligned}E(X_{n+1} \mid X_n=k) &= (k+1)P\{X_{n+1}=k+1 \mid X_n=k\}+kP\{X_{n+1}=k \mid X_n=k\} \\ &= (k+1)\frac{k}{n+2}+k\frac{(n+2-k)}{n+2} \\ &= k+\frac{k}{n+2}\end{aligned}$$

即

$$E(X_{n+1} \mid X_n)=X_n+\frac{X_n}{n+2}$$

由于M_n表示第n次抽取后的坛子中的红球所占的比例，则$M_n=\frac{X_n}{n+2}$，故M_n关于$\mathscr{F}_n=\sigma(X_1,X_2,\cdots,X_n)$可测，且

$$\begin{aligned}E(M_{n+1} \mid \mathscr{F}_n) &= E(M_{n+1} \mid X_n) \\ &= E\left(\frac{X_{n+1}}{n+1+2} \mid X_n\right) \\ &= \frac{1}{n+3}E(X_{n+1} \mid X_n) \\ &= \frac{1}{n+3}\left(X_n+\frac{X_n}{n+2}\right) \\ &= \frac{X_n}{n+2}=M_n\end{aligned}$$

6.2 鞅的停时定理

一般而言，关于$\{X_n\}$的鞅$\{M_n,n\geqslant 0\}$，易知对$\forall n\geqslant 0$，有$E(M_n)=E(M_0)$成立. 根据实际需要，通常更关心的是对于任意随机时刻T，是否仍然有$E(M_T)=E(M_0)$成立. 下面借助鞅的停时定理来讨论这个问题.

定义 6.8 设$\{X_n,n\geqslant 0\}$是一个随机变量序列，称随机函数T是关于$\{X_n,n\geqslant 0\}$的**停时**，如果T取值于$(0,1,2,\cdots,+\infty)$，且对任意n，有$\{T=n\}\in\mathscr{F}_n=\sigma(X_1,X_2,\cdots,X_n)$.

【例 6.7】 若$T\equiv n$，则T是停时，即在赌博开始时已确定n局之后一定结束.

【例 6.8】 首达时是停时. $\{X_n,n\geqslant 0\}$是一个随机变量序列，设A为一事件集

合，令

$$T(A)=\inf\{n:X_n\in A\},T(\varnothing)=\inf\{n:X_n\in\varnothing\}=+\infty$$

则 $T(A)$ 是停时．事实上

$$\{T(A)=n\}\Leftrightarrow\{X_0\notin A,X_1\notin A,\cdots,X_{n-1}\notin A,X_n\in A\}$$

显然$\{T(A)=n\}$完全由 $X_0,X_1,X_2,\cdots,X_n$ 决定，从而 $T(A)$ 是关于$\{X_n,n\geqslant 0\}$的停时．

定理6.1　设 T 为任意的取值于$(0,1,2,\cdots,+\infty)$的随机变量，则下列论断等价：

(1) T 是停时；

(2) $\{T=n\}\in\mathscr{F}_n$；

(3) $\{T\leqslant n\}\in\mathscr{F}_n$；

(4) $\{T>n\}\in\mathscr{F}_n$．

注意到以下事实：

$$\{T\leqslant n\}=\bigcup_{k=0}^{n}\{T=k\}\in\mathscr{F}_n$$

$$\{T>n\}=\Omega-\{T\leqslant n\}\in\mathscr{F}_n$$

$$\{T=n\}=\{T\leqslant n\}-\{T\leqslant n-1\}\in\mathscr{F}_n$$

故上述结论成立．

定理6.2　设 T、S 是两个停时，则 $T+S$、$\min(T,S)$、$\max(T,S)$ 都是停时．

证明

$$\{T+S>n\}=\{T+S>n,0<T<n\}\cup\{T+S>n,T=0\}$$
$$\cup\{T>n,S=0\}\cup\{T>n,S>0\}\in\mathscr{F}_n$$
$$\{\min(T,S)>n\}=\{T>n\}\cap\{S>n\}\in\mathscr{F}_n$$
$$\{\max(T,S)\leqslant n\}=\{T\leqslant n\}\cap\{S\leqslant n\}\in\mathscr{F}_n$$

定理6.3　（有界停时定理）设$(M_n,\mathscr{F}_n)$是鞅，T 是关于$\{\mathscr{F}_n,n\geqslant 0\}$的停时，若存在常数 K 使得 $T\leqslant K$，则

$$E(M_T\mid\mathscr{F}_0)=M_0$$
$$E(M_T)=E(M_0)$$

证明　由于 $T\leqslant K$，故 T 只取有限个值．当 $T=j$ 时，$M_T=M_j$，于是

$$M_T=\sum_{j=0}^{K}M_jI_{\{T=j\}}$$

$$E(M_T\mid\mathscr{F}_{K-1})=E(\sum_{j=0}^{K}M_jI_{\{T=j\}}\mid\mathscr{F}_{K-1})$$
$$=\sum_{j=0}^{K-1}M_jI_{\{T=j\}}+E(M_KI_{\{T=K\}}\mid\mathscr{F}_{K-1})$$

其中最后一个等式是由于当 $j\leqslant K-1$ 时，$M_jI_{\{T=j\}}$ 关于 $\mathscr{F}_{K-1}$ 可测．

由于 $T \leqslant K$,故 $\{T=K\}$ 与 $\{T>K-1\}$ 等价. 又由于 $\{T>K-1\} \in \mathscr{F}_{K-1}$,因此

$$E(M_K I_{\{T=K\}} \mid \mathscr{F}_{K-1}) = I_{\{T>K-1\}} E(M_K \mid \mathscr{F}_{K-1}) = I_{\{T>K-1\}} M_{K-1}$$

从而

$$\begin{aligned}
E(M_T \mid \mathscr{F}_{K-1}) &= \sum_{j=0}^{K-1} M_j I_{\{T=j\}} + E(M_K I_{\{T=K\}} \mid \mathscr{F}_{K-1}) \\
&= \sum_{j=0}^{K-1} M_j I_{\{T=j\}} + I_{\{T>K-1\}} M_{K-1} \\
&= \sum_{j=0}^{K-2} M_j I_{\{T=j\}} + M_{K-1}(I_{\{T=K-1\}} + I_{\{T>K-1\}}) \\
&= \sum_{j=0}^{K-2} M_j I_{\{T=j\}} + M_{K-1} I_{\{T\geqslant K-1\}} \\
&= \sum_{j=0}^{K-2} M_j I_{\{T=j\}} + M_{K-1} I_{\{T>K-2\}}
\end{aligned}$$

类似地,有

$$\begin{aligned}
E(M_T \mid \mathscr{F}_{K-2}) &= E[E(M_T \mid \mathscr{F}_{K-1}) \mid \mathscr{F}_{K-2}] \\
&= E(\sum_{j=0}^{K-2} M_j I_{\{T=j\}} + M_{K-1} I_{\{T>K-2\}} \mid \mathscr{F}_{K-2}) \\
&= \sum_{j=0}^{K-2} M_j I_{\{T=j\}} + I_{\{T>K-2\}} E(M_{K-1} \mid \mathscr{F}_{K-2}) \\
&= \sum_{j=0}^{K-2} M_j I_{\{T=j\}} + M_{K-2} I_{\{T>K-2\}} \\
&= I_{\{T>K-3\}} M_{K-2} + \sum_{j=0}^{K-3} M_j I_{\{T=j\}}
\end{aligned}$$

重复上述过程得

$$E(M_T \mid \mathscr{F}_0) = I_{\{T\geqslant 0\}} M_0 = M_0$$

在定理 6.3 中,严格要求满足 $T \leqslant K$,即 T 有界,从实际意义和概率意义来看,此条件都太强. 事实上,假设 T 是一个停时,且满足 $P\{T<+\infty\}=1$,也就是说以概率 1 可以保证会停止,也能在此条件下证明定理 6.3 的结论成立.

考虑停时 $T_n = \min(T, n)$,由

$$M_T = M_{T_n} + M_T I_{\{T>n\}} - M_n I_{\{T>n\}}$$

$$E(M_T) = E(M_{T_n}) + E(M_T I_{\{T>n\}}) - E(M_n I_{\{T>n\}})$$

易见,T_n 是一个有界停时,由定理 6.3 知 $E(M_{T_n}) = E(M_0)$. 由于 $P\{T<+\infty\}=1$,因此,当 $n \to +\infty$,$P\{T>n\} \to 0$. 若 $E(|M_T|)<+\infty$,则 $E(M_T I_{\{T>n\}}) \to 0 (n \to +\infty)$. 然而第三项 $E(M_n I_{\{T>n\}}) \to 0$ 却不一定成立. 考虑用例 6.5 来解释这个问题.

由例 6.5 知,$(W_n, \mathscr{F}_n)$ 是鞅,且

$$P\{W_{n+1}=1 \mid W_n=1\}=1$$

$$P\{W_{n+1}=1 \mid W_n=-(2^n-1)\}=\frac{1}{2}$$

$$P\{W_{n+1}=-2^n-2^n+1 \mid W_n=-(2^n-1)\}=\frac{1}{2}$$

设 T 表示赌博停止的时刻，则 T 是一个停时，且

$$P\{T>n\}=2^{-n}$$

$$E(M_n I_{\{T>n\}})=2^{-n}(1-2^n)$$

当 $n\to+\infty$ 时，上式并不趋于 0，这也是为什么有界停时定理的结论在此处不成立的原因.

如果要求 M_n 和 T 满足

$$\lim_{n\to+\infty} E(|M_n I_{\{T>n\}}|)=0$$

则可得到如下条件较为宽松的鞅停时定理：

定理 6.4 （鞅停时定理）设 $M_1,M_2,\cdots,M_n,\cdots$ 是关于 $\mathscr{F}_n$ 的鞅，T 是停时且满足

(1) $P\{T<+\infty\}=1$；

(2) $E(|M_T|)<+\infty$；

(3) $\lim\limits_{n\to+\infty} E(|M_n I_{\{T>n\}}|)=0$.

则有 $E(M_T)=E(M_0)$.

【例 6.9】 设 $X_0=a$，X_n 是在 $\{0,1,2,\cdots,N\}$ 上的简单随机游动 ($p=1/2$)，并且 0 和 N 是两个吸收壁. 令 $T=\min\{j:X_j=0 \text{ 或 } N\}$ 为吸收时刻，求

(1) 在吸收时刻它在 N 点的概率；

(2) 在被吸收时刻的平均时间.

解 (1) 首先证明 X_n 是一个鞅.

根据简单随机游动的定义，当 $k=1,2,\cdots,N-1$ 时

$$P\{X_{n+1}=k+1 \mid X_n=k\}=P\{X_{n+1}=k-1 \mid X_n=k\}=\frac{1}{2}$$

$$E(X_{n+1} \mid X_n=k)=k$$

当 $k=0,N$ 时，

$$P\{X_{n+1}=k \mid X_n=k\}=1$$

因此 $E(X_{n+1} \mid \mathscr{F}_n)=E(X_{n+1} \mid X_n)=X_n$，故 X_n 是一个鞅.

由于 X_n 是 Markov 链，0 和 N 是两个吸收壁，故 $P\{T<+\infty\}=1$.

由于 X_n 取值有界，所以 $E(|X_T|)<+\infty$，且

$$E(|X_n I_{\{T>n\}}|)\leqslant NE(I_{\{T>n\}})\to 0, n\to+\infty$$

故满足定理 6.4 的条件，于是

$$E(X_T) = E(X_0) = a$$

由于此时 X_T 只取两个值 N 和 0，则

$$E(X_T) = N \cdot P\{X_T = N\} + 0 \cdot P\{X_T = 0\}$$

从而得到

$$P\{X_T = N\} = \frac{E(X_T)}{N} = \frac{a}{N}$$

(2) 设 X_n 是直线上的随机游动，0 为吸收壁，已知 $X_0 = a(0 < a < N)$，定义 $M_n = X_n^2 - n$，则

$$E(M_{n+1} \mid \mathscr{F}_n) = E[X_{n+1}^2 - (n+1) \mid \mathscr{F}_n] = X_n^2 + 1 - (n+1) = X_n^2 - n = M_n$$

因此，M_n 是关于 $\mathscr{F}_n = \sigma(X_1, X_2, \cdots, X_n)$ 的鞅，但是 M_n 不是有界鞅.

设 $T = \min\{j : X_j = 0 \text{ 或 } N\}$，则 T 是一个停时，可以证明存在 $C < +\infty, 0 < \rho < 1$ 使得

$$P\{T > n\} \leqslant C\rho^n$$

因此

$$E(T) = \sum_{n=0}^{+\infty} P\{T > n\} \leqslant \sum_{n=0}^{+\infty} C\rho^n < +\infty$$

由 $M_T \leqslant N^2 + T$ 知，

$$E(M_T) \leqslant E(N^2 + T) \leqslant N^2 + E(T) < +\infty$$

并且

$$E(|M_n| I_{(T>n)}) \leqslant C\rho^n (N^2 + n) \to 0$$

从而满足停时定理 6.4 的条件，因此

$$E(M_T) = E(M_0) = a^2$$

注意到，

$$\begin{aligned} E(M_T) &= E(X_T^2) - E(T) \\ &= N^2 P\{X_T = N\} + 0 \times P\{X_T = 0\} - E(T) = aN - E(T) \end{aligned}$$

从而，

$$E(T) = aN - a^2 = a(N - a)$$

下面不加证明地给出有关上鞅的停时定理的某些结果：

定理 6.5 （上鞅的停时定理）设 $\{M_n, n \geqslant 0\}$ 是关于 $\{X_n, n \geqslant 0\}$ 的上鞅，T 是 $\{X_n, n \geqslant 0\}$ 的停时，$T_n = \min(T, n)$. 设存在一非负随机变量 W，满足 $E(W) < +\infty$，且使得

$$M_{T_n} \geqslant -W, \forall n \geqslant 0$$

则有

$$E(M_0) \geqslant E(M_T I_{\{T<+\infty\}})$$

特别地，若 $P\{T < +\infty\} = 1$，则有

$$E(M_0) \geqslant E(M_T)$$

推论6.1　设$\{M_n,n\geqslant 0\}$是关于$\{X_n,n\geqslant 0\}$的上鞅，T是$\{X_n,n\geqslant 0\}$的停时，且$M_n\geqslant 0$，则有

$$E(M_0)\geqslant E(M_T I_{\{T<n\}})$$

对于任意的$n\geqslant 0$，上鞅显然满足$E(M_0)\geqslant E(M_n)$. 定理6.5说明，当把n换成T时，在某些附加条件下，此结论也成立.

6.3　鞅的收敛定理

本节主要介绍鞅的收敛性质，即在一般的条件下，鞅$\{M_n\}$可以收敛到一个随机变量$M_{+\infty}$. 先考虑从如下的例子入手：

【例6.6(续)】　设M_n表示第n次摸球后红球所占的比例，当$n\to+\infty$时，这个比例将如何变化?

解　设$0<a<b<1$，$M_n<a$，令

$$T=\min\{j:j\geqslant n,M_j\geqslant b\}$$

即T表示n次摸球后第一次比例从小于a到超过b的时刻，则T是一个停时. 令$T_m=\min\{T,m\}$，则对$m>n$，由定理6.4可知

$$E(M_{T_m})=M_n<a$$

但是

$$E(M_{T_m})\geqslant E(M_{T_m}I_{\{T\leqslant m\}})=E(M_T I_{\{T\leqslant m\}})\geqslant bP\{T\leqslant m\}$$

从而

$$P\{T\leqslant m\}<\frac{a}{b}$$

因为上式对一切$m>n$成立，于是有

$$P\{T<+\infty\}\leqslant\frac{a}{b}$$

这说明红球比例至少以概率$1-\dfrac{a}{b}$不会超过b.

若假定这一比例确实超过b，由同样的讨论可知，红球比例再一次降到a以下的概率最大为$\dfrac{1-b}{1-a}$. 继续同样的讨论，可以知道，从a出发超过b，再小于b，…，有n个循环的概率应为

$$\left(\frac{a}{b}\right)\left(\frac{1-b}{1-a}\right)\left(\frac{a}{b}\right)\cdots\left(\frac{a}{b}\right)\left(\frac{1-b}{1-a}\right)=\left(\frac{a}{b}\right)^n\left(\frac{1-b}{1-a}\right)^n\to 0,n\to+\infty$$

因此，这个比例不会在a，b之间无限次地跳跃，a，b的任意性，也表明这一比例不会在任意的两个数之间无限次地跳跃. 也就是说，极限$\lim\limits_{n\to+\infty}M_n$存在，记为

$M_{+\infty}$，可以证明 $M_{+\infty}$ 是一个随机变量且服从$[0,1]$上的均匀分布.

下面不加证明地给出如下一般结论：

定理 6.6 (Doob 收敛定理) 设$\{M_n\}$是关于$\{\mathscr{F}_n\}$的鞅，若存在常数 $C < +\infty$，使得 $E(|M_n|) < C$ 对任意 n 成立，则当 $n \to +\infty$ 时，$\{M_n\}$几乎处处收敛到一个随机变量 $M_{+\infty}$.

【例 6.10】 设 $X_1, X_2, \cdots, X_n, \cdots$ 是独立同分布的随机变量序列，$P\{X_i = 3/2\} = P\{X_i = 1/2\} = 1/2$，令 $M_0 = 1$，对 $n > 0$，令 $M_n = X_1 \cdot X_2 \cdot \cdots \cdot X_n$，则可以证明 M_n 是关于 $X_1, X_2, \cdots, X_n, \cdots$ 的鞅. 由于 $E(|M_n|) = 1$，因此鞅收敛定理的条件成立，从而当 $n \to +\infty$ 时有

$$M_n \to M_{+\infty}$$

6.4 布朗运动的基本概念和性质

1827 年英国植物学家布朗发现液体中悬浮的花粉粒具有无规则的运动，这种运动就是布朗运动. 布朗运动(Brownian Motion)是指悬浮在流体中的微粒受到流体分子与粒子的碰撞而发生的不停息的随机运动. 然而真正用于描述布朗运动随机过程的定义是维纳(Wiener)给出的，因此布朗运动又称为维纳过程.

6.4.1 布朗运动的定义

现从讨论简单的随机游动开始，设有一质点在直线上做随机游动，质点每经过 Δt 时间，随机地以概率 $p = \frac{1}{2}$ 向右移动 $\Delta x > 0$ 的距离，以概率 $q = \frac{1}{2}$ 向左移动 $\Delta x > 0$ 的距离，且每次移动都是相互独立的.

若以 $X(t)$ 记时刻 t 质点的位置，则

$$X(t) = \Delta x (X_1 + X_2 + \cdots + X_{[t/\Delta t]}) \tag{6.1}$$

其中$[t/\Delta t]$ 表示 $t/\Delta t$ 的整数部分，令

$$X_i = \begin{cases} 1, & \text{如果第 } i \text{ 步向右} \\ -1, & \text{如果第 } i \text{ 步向左} \end{cases}$$

且假设诸 X_i 相互独立，易得

$$P\{X_i = 1\} = P\{X_i = -1\} = \frac{1}{2}$$

$$E(X_i) = 0, D(X_i) = E(X_i^2) = 1$$

$$E[X(t)] = 0, D[X(t)] = (\Delta x)^2 \left(\frac{t}{\Delta t}\right)$$

其中 $\Delta x = c\sqrt{\Delta t}$，$c > 0$ 且为常数，它由具体物理意义确定.

令 $\Delta t \to 0$，即研究连续游动，则有 $E[X(t)] = 0$，$\mathrm{Var}[X(t)] \to c^2 t$.

由式(6.1)及中心极限定理可得这一极限过程如下的一些直观性质：

(1) 当 $\Delta t \to 0$ 时，$X(t)$ 趋向于正态分布，即 $X(t) \sim N(0, c^2 t)$；

(2) $\{X(t), t \geqslant 0\}$ 有独立的增量；

(3) $\{X(t), t \geqslant 0\}$ 有平稳的增量.

由此，引出布朗(Brown)运动的定义：

定义 6.9　若一随机过程 $\{B(t); t \geqslant 0\}$ 满足：

(1) $\forall s, t > 0$，$B(s+t) - B(s) \sim N(0, c^2 t)$；

(2) $\{B(t)\}$ 是独立增量过程；

(3) $B(t)$ 是关于 t 的连续函数.

则称 $\{B(t); t \geqslant 0\}$ 为 Brown 运动，也称为 Wiener 过程，也常记为 $\{W(t); t \geqslant 0\}$.

若 $c = 1$，$B(0) = 0$，则称 $\{B(t); t \geqslant 0\}$ 为标准布朗运动. 它在 t 时刻的概率密度为

$$f_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$$

下面仅讨论标准布朗运动的某些性质.

定义 6.10　设 $\{X(t), t \geqslant 0\}$ 是随机过程，如果它的有限维分布是空间平移不变的，即

$$\begin{aligned} &P\{X(t_1 \leqslant x_1), X(t_2 \leqslant x_2), \cdots, X(t_n \leqslant x_n) \mid X(0) = 0\} \\ &\quad = P\{X(t_1) \leqslant x_1 + x, X(t_2) \leqslant x_2 + x, \cdots, X(t_n) \leqslant x_n + x \mid X(0) = x\} \end{aligned}$$

则称此过程为空间齐次的.

由上述定义知，标准布朗运动具有空间齐次性.

6.4.2　布朗运动的性质

定理 6.7　布朗运动 $\{B(t); t \geqslant 0\}$ 具有如下性质：

(1) $\{B(t); t \geqslant 0\}$ 是均方连续的；

(2) $\{B(t); t \geqslant 0\}$ 是鞅；

(3) $\{B(t); t \geqslant 0\}$ 是马尔可夫过程；

(4) $\{B(t); t \geqslant 0\}$ 是高斯过程(Gaussian process).

证明　(1) 由定义 6.8 知

$$E\{[B(t) - B(s)]^2\} = \mathrm{Var}[B(t) - B(s)] = c^2 |t - s|$$

因此，

$$\lim_{h\to 0}E\{[B(t+h)-B(t)]^2\}=\lim_{h\to 0}c^2|h|=0$$

于是$\{B(t);t\geqslant 0\}$是均方连续的.

(2) 因为布朗运动具有独立增量性,对任意的$s<t$,有

$$\begin{aligned}E[B(t)\mid B(y),y\leqslant s]&=E[B(s)+B(t)-B(s)\mid B(y),y\leqslant s]\\&=B(s)+E[B(t)-B(s)\mid B(y),y\leqslant s]\\&=B(s)+E[B(t)-B(s)]\\&=B(s)+0+0\\&=B(s)\end{aligned}$$

因此,$\{B(t);t\geqslant 0\}$是一个鞅.

(3) 任何具有独立增量性的随机过程都是一个马尔可夫过程,因此$\{B(t);t\geqslant 0\}$显然是一个马尔可夫过程.

(4) 所谓Gauss过程是指所有有限维分布都是多元正态分布的随机过程.可以证明布朗运动$\{B(t);t\geqslant 0\}$是均值函数$m(t)=0$,协方差函数$\gamma(s,t)=\min(t,s)$的Gauss过程.

由于布朗运动均值是0,所以其协方差函数为

$$\gamma(s,t)=\text{Cov}[B(t)B(s)]=E[B(t)B(s)]$$

若$t<s$,则$B(s)=B(t)+B(s)-B(t)$,且由独立增量性可得

$$E[B(t)B(s)]=E[B^2(t)]+E\{B(t)[B(s)-B(t)]\}=E[B^2(t)]=t$$

类似地,若$t>s$,则$E[B(t)B(s)]=s$.再由上述定理及数学归纳法可得$B(t)$的任何有限维分布都是正态的.

定理 6.8 设$\{B(t);t\geqslant 0\}$是标准布朗运动,则对于任意的常数$\alpha\neq 0$有

(1)$Y(t)=e^{\alpha B(t)-\alpha^2 t/2}$是鞅;

(2)$Y(t)=B^2(t)-t$是鞅.

证明 (1) 对$s<t$,有

$$\begin{aligned}E[e^{\alpha B(t)-\alpha^2 t/2}\mid B(y),y\leqslant s]&=E\{e^{\alpha[B(s)+B(t)-B(s)]-\alpha^2 t/2}\mid B(y),y\leqslant s\}\\&=e^{\alpha B(s)-\alpha^2 t/2}E\{e^{\alpha[B(t)-B(s)]}\mid B(y),y\leqslant s\}\\&=e^{\alpha B(s)-\alpha^2 t/2}E\{e^{\alpha[B(t)-B(s)]}\}\end{aligned}$$

由$B(t)\sim N(0,t)$的矩母函数$E[e^{\alpha B(t)}]=e^{t\alpha^2/2}<+\infty$得

$$E[e^{\alpha[B(t)-B(s)]}\}=e^{\frac{1}{2}\alpha^2(t-s)}$$

因此

$$E[e^{\alpha B(t)-\alpha^2 t/2}\mid B(y),y\leqslant s]=e^{\alpha B(s)-\alpha^2 s/2}$$

(2) 对$s<t$,因为$B(s)$和$B(t)-B(s)$是相互独立的,且$E[B(t)]=0$,则

$$
\begin{aligned}
&E[B^2(t)-t|B(y),y\leqslant s]\\
&=E\{[B(s)+B(t)-B(s)]^2-t|B(y),y\leqslant s\}\\
&=B^2(s)+E\{2B(s)[B(t)-B(s)]+[B(t)-B(s)]^2-t|B(y),y\leqslant s\}\\
&=B^2(s)+E\{[B(t)-B(s)]^2\}-t\\
&=B^2(s)+(t-s)-t\\
&=B^2(s)-s
\end{aligned}
$$

于是,结论成立.

6.5　常见的布朗运动的变化形式

6.5.1　等价变换

定理6.9　设$\{B(t);t\geqslant 0\}$是标准布朗运动,则下列随机过程也是标准布朗运动:

(1) $\{X(t),t\geqslant 0\}$,其中 $X(t)=cB(t/c^2),c>0$;

(2) $\{Y(t),t\geqslant 0\}$,其中 $Y(t)=B(t+h)-B(h),h>0$;

(3) $\{Z(t),t\geqslant 0\}$,其中 $Z(t)=\begin{cases}tB(1/t), & t>0\\ 0, & t<0\end{cases}$

证明　根据定义6.8来证明.很显然,随机过程(1)～(3)的起始点为$X(0)=Y(0)=Z(0)=0$.易于证明过程(1)～(3)具有独立正态分布增量性质,且均值函数都为0.只需要进一步证明随机过程(1)～(3)在任意区间$[s,t](s<t)$内的增量的方差等于$t-s$即可.

(1)

$$
\begin{aligned}
&\mathrm{Var}[X(t)-X(s)]=E\{[X(t)-X(s)]^2\}\\
&=E[X^2(t)]-2\mathrm{Cov}[X(s)-X(t)]+E[X^2(s)]\\
&=c^2\{E[B^2(t/c^2)]-2\mathrm{Cov}[B^2(s/c^2),B^2(t/c^2)]+E[B^2(s/c^2)]\}\\
&=c^2\left(\frac{t}{c^2}-2\frac{s}{c^2}+\frac{s}{c^2}\right)=t-s.
\end{aligned}
$$

(2)

$$
\begin{aligned}
&\mathrm{Var}[Y(t)-Y(s)]=E\{[B(t+h)-B(s+h)]^2\}\\
&=E[B^2(t+h)]-2\mathrm{Cov}[B(s+h),B(t+h)]+E[B^2(s+h)]\\
&=(t+h)-2(s+h)+(s+h)=t-s.
\end{aligned}
$$

(3)

$$
\begin{aligned}
&\mathrm{Var}[Z(t)-Z(s)]=E\{[tB(1/t)-sB(1/s)]^2\}\\
&=t^2E[B^2(1/t)]-2st\mathrm{Cov}[B(1/s),B(1/t)]+s^2E[B^2(1/s)]
\end{aligned}
$$

$$= t^2 \cdot \frac{1}{t} - 2st \cdot \frac{1}{t} + s^2 \cdot \frac{1}{s} = t - s.$$

因此，定理得证.

6.5.2 布朗桥

定义 6.11 设$\{B(t), t \geqslant 0\}$是布朗运动. 令

$$\bar{B}(t) = B(t) - tB(1), \quad 0 \leqslant t \leqslant 1$$

则称随机过程$\{\bar{B}(t), 0 \leqslant t \leqslant 1\}$为布朗桥(Brown bridge).

因为布朗运动是 Gauss 过程，所以布朗桥也是 Gauss 过程，其 n 维分布由均值函数和方差函数完全确定，且对任何 $0 \leqslant s \leqslant t \leqslant 1$，有

$$E[\bar{B}(t)] = 0$$

$$\begin{aligned} E[\bar{B}(s)\bar{B}(t)] &= E\{[B(s) - sB(1)][B(t) - tB(1)]\} \\ &= E[B(s)B(t) - tB(s)B(1) - sB(t)B(1) + tsB^2(1)] \\ &= s - ts - ts + ts = s(1-t) \end{aligned}$$

此外，由定义 6.8 可知 $\bar{B}(0) = \bar{B}(1) = 0$，即此过程的起始点是固定的，就像桥一样(图 6.1)，这就是 Brown 桥名称的由来.

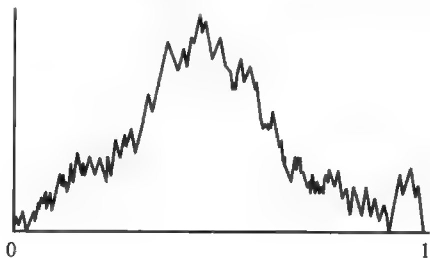


图 6.1

6.5.3 反射布朗运动

定理 6.10 如果随机过程$\{X(t), t \geqslant 0\}$满足$X(t) = |B(t)|$，就称之为反射布朗运动(在 t 轴反射出来). 它的均值和方差函数分别为

$$m(t) = E[X(t)] = \frac{2}{\sqrt{2\pi t}\sigma}\int_0^{+\infty} x e^{-\frac{x^2}{2\sigma^2 t}} dx = \sigma\sqrt{\frac{2t}{\pi}}, t \geqslant 0$$

$$\operatorname{Var}[X(t)] = E[X^2(t)] - \{E[X(t)]\}^2 = \sigma^2 t - \sigma^2 \frac{2t}{\pi} = (1 - \frac{2}{\pi})\sigma^2 t$$

事实上，反射布朗运动可看作是一齐次马尔可夫过程，状态空间 $Z = [0, +\infty)$，可以写为

$$0 \leqslant t_1 < t_2 < \cdots < t_n < +\infty, \quad x_i \in \mathbf{Z}$$

考虑到布朗运动的马尔可夫性和它关于 t 轴对称的随机性,有

$$\begin{aligned} & P\{X(t) \leqslant y \mid X(t_1) = x_1, X(t_2) = x_2, \cdots, X(t_n) = x_n\} \\ = & P\{-y \leqslant B(t) \leqslant +y \mid B(t_1) = \pm x_1, B(t_2) = \pm x_2, \cdots, B(t_n) = \pm x_n\} \\ = & P\{-y \leqslant B(t) \leqslant +y \mid B(t_n) = \pm x_n\} \end{aligned}$$

因此,当 $0 \leqslant s < t$ 时,转移概率为

$$P\{X(t) \leqslant y \mid X(s) = x\}$$

此反射布朗运动的概率由布朗运动在 $[s,t]$ 内的增量所决定,假定起始点为状态 x 和时间 s, 因为这样的增量服从 $N(x,\sigma^2\tau)$,其中 $\tau = t - s$,所以

$$P\{X(t) \leqslant y \mid X(s) = x\} = \int_{-y}^{y} e^{-\frac{(u-x)^2}{2\sigma^2\tau}} du$$

相当于

$$P\{X(t) \leqslant y \mid X(s) = x\} = \Phi(\frac{y-x}{\sigma\sqrt{\tau}}) - \Phi\left(-\frac{y+x}{\sigma\sqrt{\tau}}\right); \quad x,y \geqslant 0; \quad \tau = t - s$$

因为转移概率仅仅通过 $\tau = t - s$ 来体现 t 和 s 之间的关系,所以反射布朗运动也是齐次马尔可夫过程.

6.5.4 几何布朗运动

由 $X(t) = e^{B(t)}, t \geqslant 0$ 定义的过程 $\{X(t), t \geqslant 0\}$ 称为几何布朗运动. 由于布朗运动的矩母函数为 $E[e^{sB(t)}] = e^{s^2/2}$,所以几何布朗运动的均值函数与方差函数分别为

$$E[X(t)] = E[e^{B(t)}] = e^{1/2}$$

$$\begin{aligned} \operatorname{Var}[X(t)] &= E[X^2(t)] - \{E[X(t)]\}^2 \\ &= E[e^{2B(t)}] - e = e^{2t} - e \end{aligned}$$

早在 1900 年,法国数学家巴舍利耶(L. Bachelier)就在其博士论文《投资理论》中,给出了布朗运动的数学描述,并提出用几何布朗运动来模拟股票价格的变化. 在金融市场中,经常假定股票的价格按照几何布朗运动变化,在下面的例子中可以如此假定.

【例 6.11】 股票期权的价格.

设某人拥有某种股票交割时刻为 T,交割价格为 K 的欧式看涨期权,即他(她)具有在时刻 T 以固定的价格 K 购买一股这种股票的权利. 假定这种股票目前的价格为 y,并按照几何 Brown 运动变化,计算这个股票期权的平均价格.

解 设 $X(T)$ 表示时刻 T 的股票价格,若 $X(T)$ 高于 K,期权将被实施,因此该期权在时刻 T 的平均价格应为

$$
\begin{aligned}
E\{\max[X(T)-K,0]\} &= \int_0^{+\infty} P\{X(T)-K>u\}\mathrm{d}u \\
&= \int_0^{+\infty} P\{y\mathrm{e}^{B(T)}-K>u\}\mathrm{d}u \\
&= \int_0^{+\infty} P\left\{B(T)>\ln\frac{K+u}{y}\right\} \\
&= \frac{1}{\sqrt{2\pi T}}\int_0^{+\infty}\int_{\ln[(K+u)/y]}^{+\infty}\mathrm{e}^{-x^2/2T}\mathrm{d}x\mathrm{d}u
\end{aligned}
$$

6.5.5 Ornstein-Uhlenbeck 过程

如果布朗运动过程对于描述液体和气体中微粒的运动是完全正确的，那么微粒是以一种极大的速度在运动. 为了克服这种非现实的假设情况，学者 Ornstein 和 Uhlenbeck 建立了新的随机过程来描述这种微粒的高速运动.

定理 6.11 设$\{B(t),t\geqslant 0\}$是布朗运动，参数为σ，随机过程$\{U(t),t\in(-\infty,+\infty)\}$由如下表达式定义：

$$U(t)=\mathrm{e}^{-\alpha t}B(\mathrm{e}^{2\alpha t})$$

称之为参数为σ和$\alpha(\alpha>0)$的 Ornstein-Uhlenbeck 过程，简称 O-U 过程.

$U(t)$的概率密度函数为

$$f_{U(t)}(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-x^2/(2\sigma^2)},\quad -\infty<x<+\infty$$

因此，$U(t)$是正态分布，且

$$E[U(t)]=0,\quad \mathrm{Var}[U(t)]=\sigma^2$$

6.5.6 带漂移的布朗运动

定义 6.12 一个随机过程被称为带漂移的布朗运动，如果它满足以下条件：

(1) $D(0)=0$；

(2) $\{D(t),t\geqslant 0\}$是齐次独立增量的；

(3) 增量$D(t)-D(s)$的均值为$\mu(t-s)$，方差为$\sigma^2|t-s|$.

漂移布朗运动的等价定义为：$\{D(t),t\geqslant 0\}$是一个漂移布朗运动当且仅当$D(t)$的结构为

$$D(t)=\mu t+B(t)$$

其中$\{B(t),t\geqslant 0\}$是布朗运动，常数μ为漂移系数或简单飘移. 因此，漂移布朗运动是由布朗运动加上一个特定的函数组成的. 这个特定的函数是条直线，且正好和漂

移布朗运动的均值函数吻合,即

$$m(t) = E[D(t)] = \mu t$$

如果条件(2)和(3)都被满足,过程从 $t=0$ 和 $u(u \neq 0)$ 开始,所得到的随机过程 $\{D_u(t), t \geqslant 0\}$ 就叫作平移漂移布朗运动. $D_u(t)$ 的结构为

$$D_u(t) = u + D(t)$$

一维漂移布朗运动的密度函数为

$$f_{D(t)}(x) = \frac{1}{\sqrt{2\pi t}\sigma} e^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}; \quad -\infty < x < +\infty, \quad t > 0$$

漂移布朗运动有着广泛的应用背景,比如,在研究计算磨损参数、控制成本、生产标准和资本增量、模拟物理噪音等方面. 总体而言,漂移布朗运动可以应用于任意线性过程中对随机运动产生持久干扰的情况.

【例 6.12】 实施股票期权.

假设某人有在将来某个时刻以固定价格 A 购买一股股票的期权,与现在的市价无关,不妨取现在的市价为 0,并假定其变化遵循具有负漂移系数 $-\mu(\mu > 0)$ 的布朗运动. 问在什么时候实施期权?

解 考虑在市价为 x 时实施期权的策略. 在此策略下的平均所得是

$$(x - A)P(x)$$

其中 $P(x)$ 是过程迟早到达 x 的概率. 于是有

$$P(x) = e^{-2\mu x}, \quad x > 0$$

x 的最优值是使 $(x - A)e^{-2\mu x}$ 最大的值,易见它是 $x = A + \frac{1}{2\mu}$.

习　题　6

1. 设 $Y_0, Y_1, \cdots$ 为一系列独立的随机变量,服从同样的分布 $N(0,1)$,试证明:$\{X_n = \sum_{i=0}^{n} Y_i^2, n = 0,1,\cdots\}$ 为离散时间鞅.
2. $Y_0, Y_1, \cdots$ 是一系列独立的随机变量,它们的均值是 $E(Y_i)$. 试证明:$\{X_n = \sum_{i=0}^{n} [Y_i - E(Y_i)], n \geqslant 0\}$ 为离散时间鞅.
3. 证明:具有独立增量,均值为常数且满足 $E[|X(t)|] < +\infty, t \in T$ 的随机过程 $\{X(t), t \in T\}$ 是一个鞅.
4. 设 L 为一个随机过程 $\{X(t), t \in T\}$ 中离散和连续时间中的停时,z 是一个正常数. 证明:$L \wedge z = \min(L, z)$ 对于 $\{X(t), t \in T\}$ 是一个停时.
5. 设 $\{B(t), t \geqslant 0\}$ 为标准布朗运动,给定 $B(s), 0 \leqslant s < t$,试计算 $B(t)$ 的条件概率密度.

6. 设 $\{B_1(t), t \geqslant 0\}$，$\{B_2(t), t \geqslant 0\}$ 为相互独立的标准布朗运动，试证明 $\{X(t) = B_1(t) - B_2(t)\}$ 是布朗运动.

7. 设 $\{\bar{B}(t), 0 \leqslant t \leqslant 1\}$ 是布朗桥，证明如下定义的随机过程 $\{Z(t), t \geqslant 0\}$

$$Z(t) = (t+1)\bar{B}\left(\frac{t}{t+1}\right)$$

是标准布朗运动.

8. 设 $\{B(t), t \geqslant 0\}$ 为标准布朗运动，求 $B(1) + B(2) + \cdots + B(n)$ 的分布，并验证 $\left\{X(t) = tB\left(\frac{1}{t}\right)\right\}$ 仍为 $[0, +\infty)$ 上的布朗运动.

9. 设 $\{B(t), t \geqslant 0\}$ 为标准布朗运动，验证 $\left\{X(t) = (1-t)B\left(\frac{t}{1-t}\right), 0 \leqslant t \leqslant 1\right\}$ 是布朗桥.

第 7 章　随机分析

在数学分析的微积分内容中，连续性、导数和积分等概念都建立在极限概念的基础上. 关于随机过程分析的研究，也需要建立相应的随机过程的连续性、导数和积分等分析概念，同样这些概念也都是建立在随机序列极限的基础上，这一部分内容统称为随机过程分析，简称随机分析.

7.1　二阶矩过程与均方极限

7.1.1　二阶矩过程

定义 7.1　如果随机过程 $\{X(t), t \in T\}$，对任意 $t \in T$，有

$$m(t) = E[X(t)] < +\infty, D(t) = E\{[X(t) - m(t)]^2\} < +\infty$$

则称该过程是**二阶矩过程**. 二阶矩过程的协方差函数总存在，且为

$$\mathrm{Cov}[X(t_1), X(t_2)] = E\{E[X(t_1) - m(t_1)][\overline{X(t_2) - m(t_2)}]\}$$

进而其他相关函数也总存在.

二阶矩过程的均值函数和协方差函数一定存在，正态随机过程是二阶矩过程. 一般地，随机过程的概率性质是由其分布函数族完全确定的，但在实际问题中很难确定出分布函数族. 正态随机过程的一、二阶矩可以完全确定其有限维分布. 因此，在很多实际问题中，只需要通过对随机过程二阶矩的讨论就足以了解随机过程的某些重要统计特征. 为简便起见，设二阶矩过程是复二阶矩过程，且其均值为零.

二阶矩过程的相关函数具有如下性质：

(1) 设 $\{X(t), t \in T\}$ 是二阶矩过程，则相关函数

$$R_X(t_2, t_1) = \overline{R_X(t_1, t_2)}$$

若 $X(t)$ 是实二阶矩过程，则 $R_X(t_1, t_2) = R_X(t_2, t_1)$，即相关函数是对称的.

(2) 二阶矩过程的相关函数 $R_X(t_2, t_1)$ 具有非负定性，即对于任意有限个 t_1，$t_2, \cdots, t_n \in T$ 和 n 个复数 $\lambda_1, \lambda_2, \cdots, \lambda_n$，有

$$\sum_{k=1}^{n}\sum_{m=1}^{n}R_X(t_k,t_m)\lambda_k\bar{\lambda}_m \geqslant 0$$

其中 n 为任意正整数.

【例 7.1】 余弦波过程 $X(t)=A\cos(\omega t+\theta), t\geqslant 0$，振幅 A 与角频率 ω 取常数，相位 $\theta\sim U(-\pi,\pi)$. 因为

$$E[X(t)]^2=\int_{-\pi}^{\pi}A^2\cos^2(\omega t+\theta)\frac{1}{2\pi}\mathrm{d}\theta\leqslant A^2, \forall t\geqslant 0$$

故 $X(t)$ 是一个二阶矩过程，且

$$E[X(t)]=\int_{-\pi}^{\pi}A\cos(\omega t+\theta)\frac{1}{2\pi}\mathrm{d}\theta=0$$

$$R_X(s,t)=E[X(t)X(s)]=\int_{-\pi}^{\pi}A^2\cos(\omega t+\theta)\cos(\omega s+\theta)\frac{1}{2\pi}\mathrm{d}\theta$$

$$=\frac{1}{2}A^2\cos[\omega(s-t)]=\frac{1}{2}A^2\cos(\omega\tau)\quad(\tau=s-t)$$

7.1.2 均方极限

对于概率空间上的随机序列 $\{X_n\}$，每个试验结果 e 都对应一个序列

$$X_1(e), X_2(e), \cdots, X_n(e), \cdots$$

若该序列对每个 e 都收敛，则称随机序列 $\{X_n\}$ 处处收敛，即 $\lim\limits_{n\to+\infty}X_n=X$，其中 X 为随机变量. 在实际问题中，这种收敛性是很难达到的. 因此，下面介绍几种在概率意义下的随机序列的收敛定义，它们不要求一定满足该序列对每个 e 都收敛的条件.

定义 7.2 如果使 $\lim\limits_{n\to+\infty}X_n(e)=X(e)$ 成立的 e 的集合的概率为 1，即

$$P\{e: \lim_{n\to+\infty}X_n(e)=X(e)\}=1$$

则称二阶矩随机序列 $\{X_n(e)\}$ 以概率 1 收敛于二阶矩随机变量 $X(e)$ 或称 $\{X_n(e)\}$ 几乎处处收敛于 $X(e)$，记作 $X_n\xrightarrow{a.e.}X$.

定义 7.3 如果对于任何 $\varepsilon>0$ 有

$$\lim_{n\to+\infty}P\{|X_n(e)-X(e)|\geqslant\varepsilon\}=0$$

则称二阶矩随机序列 $\{X_n(e)\}$ 依概率收敛于二阶矩随机变量 $X(e)$，记作 $X_n\xrightarrow{P}X$.

定义 7.4 设有二阶矩随机序列 $\{X_n\}$ 和二阶矩随机变量 X，如果有

$$\lim_{n\to+\infty}E(|X_n-X|^2)=0$$

成立，则称 $\{X_n\}$ 均方收敛于 X，记作 $X_n\xrightarrow{m.s.}X$，或者记作 $\underset{n\to+\infty}{\mathrm{l.i.m}}X_n=X$ 或 $\mathrm{l.i.m}X_n=X$（l.i.m 是 limit in mean 的缩写）.

在以上三种收敛定义中，均方收敛是最简单的收敛形式，它只涉及单独一个序列，且均方收敛可以推导出依概率收敛. 以下讨论的随机序列的收敛性，都是指均方收敛.

定理 7.1 （柯西均方收敛准则）设有二阶矩随机序列$\{X_n\}$和二阶矩随机变量X，则$\{X_n\}$均方收敛于X的充要条件是

$$\lim_{n,m\to+\infty} E(|X_n - X_m|^2) = 0$$

证明略.

定理 7.2 设$\{X_n\}$，$\{Y_n\}$都是二阶矩随机序列，U为二阶矩随机变量，$\{c_n\}$为常数序列，a,b,c为常数. 令$\mathrm{l.i.m}X_n = X$，$\mathrm{l.i.m}Y_n = Y$，$\lim\limits_{n\to+\infty} c_n = c$. 则

(1)$\mathrm{l.i.m}c_n = \lim\limits_{n\to+\infty} c_n = c$；

(2)$\mathrm{l.i.m}U = U$；

(3)$\mathrm{l.i.m}(c_nU) = cU$；

(4)$\mathrm{l.i.m}(aX_n + bY_n) = aX + bY$；

(5) $\lim\limits_{n\to+\infty} E(X_n) = E(X) = E(\mathrm{l.i.m}X_n)$；

(6) $\lim\limits_{n,m\to+\infty} E(X_n\overline{Y}_m) = E(X\overline{Y}) = E[(\mathrm{l.i.m}X_n)(\mathrm{l.i.m}\overline{Y}_n)]$，特别地，有

$$\lim_{n\to+\infty} E(|X_n|^2) = E(|X|^2) = E(|\mathrm{l.i.m}X_n|^2)$$

证明 (1)，(2)，(3) 由均方收敛定义可以得证. 下证(4)：

当$n\to+\infty$时，有

$$\begin{aligned}&E[|(aX_n + bY_n) - (aX + bY)|^2]\\ =\ &E[|a(X_n - X) + b(Y_n - Y)|^2]\\ \leqslant\ &2a^2E(|X_n - X|^2) + 2b^2E(|Y_n - Y|^2) \to 0\end{aligned}$$

由均方收敛的定义知(4) 得证.

(5) 由许瓦兹不等式即可证明.

(6) 由许瓦兹不等式有

$$\begin{aligned}&|E(X_n\overline{Y}_m) - E(X\overline{Y})| = |E(X_n\overline{Y}_m - X\overline{Y})|\\ =\ &|E[(X_n - X)(\overline{Y}_m - \overline{Y}) + X_n\overline{Y} + X\overline{Y}_m - 2X\overline{Y}]|\\ =\ &|E[(X_n - X)(\overline{Y_m - Y})] + E[(X_n - X)\overline{Y}] + E[(\overline{Y}_m - \overline{Y})X]|\\ \leqslant\ &|E[(X_n - X)(\overline{Y_m - Y})]| + |E[(X_n - X)\overline{Y}]| + |E[(\overline{Y}_m - \overline{Y})X]|\\ \leqslant\ &\sqrt{E(|X_n - X|^2)E(|Y_m - Y|^2)} + \sqrt{E(|X_n - X|^2)E(|Y|^2)}\\ &+ \sqrt{E(|Y_m - Y|^2)E(|X|^2)} \to 0 \quad (\text{当 } n,m\to+\infty \text{ 时})\end{aligned}$$

故结论成立.

定理 7.3 （洛弗准则）设有二阶矩随机序列$\{X_n\}$和二阶矩随机变量X，则$\{X_n\}$均方收敛于X的充要条件是

$$\lim_{n,m\to+\infty} E(X_n\overline{X}_m)=C(\text{常数})$$

证明 充分性：

由
$$E(|X_n-X_m|^2)=E(|X_n|^2-2X_n\overline{X}_m+|X_m|^2)$$
$$=E(|X_n|^2)-2E(|X_n\overline{X}_m|)+E(|X_m|^2)$$

因为$\lim\limits_{n,m\to+\infty} E(X_n\overline{X}_m)=C$，所以

$$\lim_{n,m\to+\infty} E(|X_n-X_m|^2)=C-2C+C=0$$

由柯西均方收敛准则知，$\{X_n\}$均方收敛于X.

必要性的证明由定理 7.2 的(6) 知成立.

【例 7.2】 设$\{X_n\}$是一随机变量序列，X是随机变量，若$\forall\varepsilon>0$，当$n\to+\infty$时，$\{X_n\}$满足$P\{|X_n-X|>\varepsilon\}\to0$，则称$X_n$依概率收敛于$X$. 证明：若$\text{l.i.m}X_n=X$，则$X_n\xrightarrow{P}X$.

证明 若$\text{l.i.m}X_n=X$，则$\lim\limits_{n\to+\infty}E(|X_n-X|^2)=0$，由契比雪夫不等式

$$P\{|X_n-X|>\varepsilon\}\leqslant E(|X_n-X|^2)/\varepsilon^2$$

可知，当$n\to+\infty$时，$E(|X_n-X|^2)\to0$，则$\forall\varepsilon>0$，必有$P\{|X_n-X|>\varepsilon\}\to0$，所以，$X_n$均方收敛于$X$，则必依概率收敛于$X$.

【例 7.3】 证明均方极限的唯一性.

证明 设$\{X(n)\}$是二阶矩随机序列，X,Y是二阶矩变量，且$\lim\limits_{n\to+\infty}X(n)=X$，$\lim\limits_{n\to+\infty}X(n)=Y$，因为

$$E\{[X(n)-Y][\overline{X(n)-Y}]\}$$
$$=E[X(n)\overline{X(n)}]-E[Y\overline{X(n)}]-E[X(n)\overline{Y}]+E(Y\overline{Y})$$

对上式两边取极限$n\to+\infty$，左边用$X(n)\to X$代入，右边用$X(n)\to Y$代入，则左边为

$$E\{[X(n)-Y][\overline{X(n)-Y}]\xrightarrow{n\to+\infty}E[(X-Y)\overline{(X-Y)}]\}=E(|X-Y|^2)$$

右边为

$$E[X(n)\overline{X(n)}]-E[Y\overline{X(n)}]-E[X(n)\overline{Y}]+E(Y\overline{Y})$$
$$\xrightarrow{n\to+\infty}E(Y\overline{Y})-E(Y\overline{Y})-E(Y\overline{Y})+E(Y\overline{Y})=0$$

所以$E(|X-Y|^2)=0$，即$X=Y$，故均方极限唯一.

【例 7.4】 设$Y_1,Y_2,\cdots$是独立同分布的随机变量序列，$E(Y_1)=\mu$，$D(Y_1)=\sigma^2$，令$X(n)=\dfrac{1}{n}\sum\limits_{i=1}^{n}Y_i$，证明：$\text{l.i.m}\,X_n=\mu$.

证明　因为

$$E[|X(n)-\mu|^2]=E\left(\left|\frac{1}{n}\sum_{i=1}^{n}Y_i-\mu\right|^2\right)=\frac{1}{n^2}E\left(\left|\sum_{i=1}^{n}(Y_i-\mu)\right|^2\right)$$

$$=\frac{1}{n}E(|Y_i-\mu|^2)$$

故

$$\lim_{n\to+\infty}E[|X(n)-\mu|^2]=\lim_{n\to+\infty}\frac{1}{n}E(|Y_i-\mu|^2)=0$$

所以
$$\text{l.i.m}X_n=\mu$$

此例又称为均方极限下的大数定律.

【例 7.5】　设$\{X_n,n\geqslant 1\}$是 Poisson 随机变量序列,证明:该序列的均方极限服从 Poisson 分布.

证明　记 $E(X)=\lambda,E(X_n)=\lambda_n$,因为 $\lim\limits_{n\to+\infty}X_n=X$,所以

$$\lim_{n\to+\infty}E(X_n)=E(\mathop{\text{l.i.m}}_{n\to+\infty}X_n)=E(X)$$

即
$$\lim_{n\to+\infty}\lambda_n=\lambda$$

设 $\varphi_n(t)$ 和 $\varphi(t)$ 分别表示 X_n 和 X 的特征函数,则

$$\lim_{n\to+\infty}\varphi_n(t)=\lim_{n\to+\infty}E(\mathrm{e}^{\mathrm{i}tX_n})$$

$$=\lim_{n\to+\infty}\mathrm{e}^{\lambda_n(\mathrm{e}^{\mathrm{i}t}-1)}=\mathrm{e}^{\lim\limits_{n\to+\infty}\lambda_n(\mathrm{e}^{\mathrm{i}t}-1)}=\mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}t}-1)}=\varphi(t)$$

因为特征函数与分布函数一一对应,且 $\varphi(t)$ 是 Poisson 分布随机变量的特征函数,故 X 服从 Poisson 分布.

7.2　均方连续与均方导数

7.2.1　随机过程的均方连续

设$\{X(t),t\in T\}$为二阶矩过程,参数集 T 为直线上的一个有限区间或者无限区间.下面讨论二阶矩过程的均方连续的性质.

定义 7.5　设$\{X_t,t\in T\}$为二阶矩过程,若对某确定的$t\in T$,有 $\text{l.i.m}X(t+h)=X(t)$ 即

$$\lim_{h\to 0}E[|X(t+h)-X(t)|^2]=0,\quad t,t+h\in T$$

则称二阶矩过程$\{X(t),t\in T\}$在点t处均方连续.若$\{X(t),t\in T\}$在 $\forall t\in T$ 处都均方连续,则称$\{X(t),t\in T\}$在 T 上均方连续.

定理7.4　（均方连续准则）设$\{X(t),t\in T\}$为二阶矩过程，$R_X(t_1,t_2)$为其相关函数，$t_1,t_2\in T$，则过程$\{X(t),t\in T\}$在t处均方连续的充分必要条件为$R_X(t_1,t_2)$在(t,t)处连续．

证明　充分性：若$R_X(t_1,t_2)$在(t,t)处连续，由

$$\begin{aligned}&\lim_{h\to0}E[|X(t+h)-X(t)|^2]\\=&\lim_{h\to0}[R_X(t+h,t+h)-R_X(t,t+h)-R_X(t+h,t)+R_X(t,t)]\end{aligned}$$

则
$$\lim_{h\to0}E[|X(t+h)-X(t)|^2]=0,t,t+h\in T$$

必要性：如果$\lim\limits_{h\to0}E[|X(t+h)-X(t)|^2]=0,t,t+h\in T$，则由定理7.2的(6)得

$$\lim_{\substack{t_1\to+\infty\\t_2\to+\infty}}R_X(t_1,t_2)=\lim_{\substack{t_1\to+\infty\\t_2\to+\infty}}E[X(t_1)\overline{X(t_2)}]=E[X(t)\overline{X(t)}]=R_X(t,t)$$

均方连续准则揭示了二阶矩过程的均方连续性可由其相关函数在普通意义下的连续性来确定．

【例7.6】　设$\{X(t),t\geqslant0\}$，$X(0)=0$是Poisson过程，讨论其均方连续性．

解　因为$\{X(t),t\geqslant0\}$是Poisson过程，所以有

$$P\{X(t)-X(s)=k\}=\mathrm{e}^{-\lambda(t-s)}\frac{[\lambda(t-s)]^k}{k!},k=0,1,\cdots,t>s$$

令$s=0$，则有$E[X(t)]=\lambda t$．

当$s<t$时，有

$$\begin{aligned}C_X(s,t)&=E\{[X(s)-\lambda s][X(t)-\lambda t]\}\\&=E\{[X(s)-\lambda s]\{[X(s)-\lambda s]+[X(t)-\lambda t]-[X(s)-\lambda s]\}\}\\&=E\{[X(s)-\lambda s]^2\}=D[X(s)]^2=\lambda s\end{aligned}$$

当$s>t$时，类似可得$C_X(s,t)=\lambda t$，故$C_X(s,t)=\lambda\min\{s,t\}$，则

$$R_X(s,t)=C_X(s,t)+m(s)m(t)=\lambda\min\{s,t\}+\lambda^2st$$

由于$C_X(s,t)$[或$R_X(s,t)$]在$\{(t,t),t\geqslant0\}$处二元连续，所以，Poisson过程$\{X(t),t\geqslant0\}$在$t\geqslant0$时均方连续．

【例7.7】　设有二阶矩过程$\{X(t),t\in T\}$，$R_X(s,t)$是相关函数．

证明：若$R_X(s,t)$在(t,t)处连续，则它在$T\times T$上连续．

证明　若$R_X(s,t)$在(t,t)处连续，则$X(t)$在t处均方连续，所以，对$s,t,s+h,t+h_1\in T$有

$$\mathrm{l.i.m}\,X(s+h)=X(s),\ \mathrm{l.i.m}\,X(t+h_1)=X(t)$$

根据二阶矩过程的均方收敛性质得

$$\begin{aligned}\lim_{h,h_1\to0}R_X(s+h,t+h_1)&=\lim_{h,h_1\to0}E[X(s+h)\overline{X(t+h_1)}]\\&=E[X(s)\overline{X(t)}]=R_X(s,t)\end{aligned}$$

即 $R_X(s,t)$ 在 $T\times T$ 上连续.

7.2.2　随机过程的均方导数

定义 7.6　设 $\{X(t),t\in T\}$ 为二阶矩过程，参数集 T 为直线上的一个有限区间或者无限区间. 若存在另一二阶矩过程 $\{X'(t),t\in T\}$，使得

$$\lim_{h\to 0}E\left[\frac{X(t+h)-X(t)}{h}-X'(t)\right]^2=0$$

即 $\underset{h\to 0}{\text{l.i.m}}\dfrac{X(t+h)-X(t)}{h}=X'(t)$，则称 $X(t)$ 在点 t 处均方可微，并称 $X'(t)$ 为 $X(t)$ 在点 t 处的均方导数，或记作 $\dfrac{\mathrm{d}X(t)}{\mathrm{d}t}$.

若 $X(t)$ 在 $(-\infty,+\infty)$ 上每一点处都均方可微，则称 $X(t)$ 在 $T\in(-\infty,+\infty)$ 上均方可微. 此时 $X'(t)$ 也是一个随机过程，与 $X(t)$ 在同一概率空间上.

类似地，可定义 $\{X'(t),t\in T\}$ 的均方导数过程：

$$\{X''(t),t\in T\},\{X'''(t),t\in T\},\cdots$$

将随机过程的均方导数转移到实数域进行讨论分析，引进广义二阶导数概念：

定义 7.7　称二元函数 $f(s,t)$ 在 (s,t) 处广义二阶可微，若极限

$$\lim_{\substack{\Delta s\to 0\\ \Delta t\to 0}}\frac{f(s+\Delta s,t+\Delta t)-f(s+\Delta s,t)-f(s,t+\Delta t)+f(s,t)}{\Delta t\Delta s}$$

存在，称此极限为 $f(s,t)$ 在 (s,t) 处的广义二阶导数.

广义二阶导数是二重极限，而二阶混合偏导是二次极限，一般情况下两者不相等. 不加证明地给出如下的广义二阶导数存在的充分条件：

定理 7.5　（广义二阶导数存在的充分条件）若二元函数 $f(s,t)$ 关于 s,t 的一阶偏导存在，二阶混合偏导存在并连续，则 $f(s,t)$ 一定是广义二阶可导的，且广义二阶导数等于二阶混合偏导数，即 $f''_{st}(s,t)=f''_{ts}(s,t)$.

定理 7.6　（均方可微准则）设 $\{X(t),t\in T=(-\infty,+\infty)\}$ 为二阶矩过程，$R_X(s,t)$ 为其相关函数，则过程 $X(t)$ 在 t_0 处均方可微的充分必要条件为 $R_X(s,t)$ 在 (t_0,t_0) 处的广义二阶导数存在.

证明　由均方收敛定义及准则可知，$\{X(t),t\in T\}$ 在 $t=t_0$ 处均方可微

$\Leftrightarrow \underset{\Delta t\to 0}{\text{l.i.m}}\dfrac{X(t_0+\Delta t)-X(t_0)}{\Delta t}$ 存在，

$\Leftrightarrow \lim\limits_{\substack{\Delta t\to 0\\ \Delta s\to 0}}E\left\{\dfrac{X(t_0+\Delta t)-X(t_0)}{\Delta t}\cdot\dfrac{X(t_0+\Delta s)-X(t_0)}{\Delta s}\right\}$ 存在，

$$\Leftrightarrow \lim_{\substack{\Delta t\to 0\\ \Delta s\to 0}} \frac{1}{\Delta t\Delta s}[R_X(t_0+\Delta t,t_0+\Delta s)-R_X(t_0+\Delta t,t_0)-R_X(t_0,t_0+\Delta s)+R_X(t_0,$$

$t_0)]$ 存在，即 $R_X(s,t)$ 在 (t_0,t_0) 处广义二阶可微.

推论 7.1 二阶矩过程 $\{X_t,t\in T\}$ 的相关函数 $R_X(s,t)$ 对任意的 $t\in T$ 在 (t,t) 处广义二阶可微，则 $R'_t(s,t)R'_t(s,t)$，$R''_{st}(s,t)$，$R''_{ts}(s,t)$ 在 $T\times T$ 上均存在，而且

(1) 导数过程 $\{X'_t,t\in T\}$ 的均值函数为

$$m_{X'}(t)=E(X'_t)=\frac{\mathrm{d}}{\mathrm{d}t}E(X_t)=m'_X(t)$$

(2) 导数过程 $\{X'_t,t\in T\}$ 的自相关函数为

$$R_{X'}(s,t)=E(X'_sX'_t)=R''_{st}(s,t)=R''_{ts}(s,t)$$

(3) 导数过程 $\{X'_t,t\in T\}$ 与 $\{X_t,t\in T\}$ 的互相关函数为

$$R_{X'X}(s,t)=E(X'_sX_t)=R'_s(s,t)$$

$$R_{XX'}(s,t)=E(X_sX'_t)=R'_t(s,t)$$

证明 (1) $m_{X'}(t)=E(X'_t)=E\left[\underset{\Delta t\to 0}{\mathrm{l.i.m}}\dfrac{X(t+\Delta t)-X(t)}{\Delta t}\right]$

$$=\lim_{\Delta t\to 0}E\left[\frac{X(t+\Delta t)-X(t)}{\Delta t}\right]$$

$$=\lim_{\Delta t\to 0}\left[\frac{m_X(t+\Delta t)-m_X(t)}{\Delta t}\right]=m'_X(t)$$

(2) $R_{X'}(s,t)=E(X'_sX'_t)=E\left[\underset{\Delta s\to 0}{\mathrm{l.i.m}}\dfrac{X(s+\Delta s)-X(s)}{\Delta s}\cdot X'(t)\right]$

$$=\lim_{\Delta s\to 0}E\left[\frac{X(s+\Delta s)X'(t)-X(s)X'(t)}{\Delta s}\right]$$

$$=\lim_{\Delta s\to 0}\frac{R'_t(s+\Delta s,t)-R'_t(s,t)}{\Delta s}=R''_{ts}(s,t)$$

(3) $R_{X'X}(s,t)=E(X'_sX_t)=E\left[\underset{\Delta s\to 0}{\mathrm{l.i.m}}\dfrac{X(s+\Delta s)-X(s)}{\Delta s}\cdot X(t)\right]$

$$=\lim_{\Delta s\to 0}E\left[\frac{X(s+\Delta s)X(t)-X(s)X(t)}{\Delta s}\right]$$

$$=\lim_{\Delta s\to 0}\frac{R_X(s+\Delta s,t)-R_X(s,t)}{\Delta s}=R'_s(s,t)$$

类似地证明第二个等式成立.

定理 7.7 均方导数也有类似于普通函数导数的性质，不加证明地归纳如下：

(1) 均方导数具有唯一性.

(2) 任一随机变量 X(或为常数) 的均方导数为零.

(3) 若 $\{X(t),t\in T\}$、$\{Y(t),t\in T\}$ 是两个均方可微的随机过程，a，b 为两个常数，则 $aX(t)+bY(t)$ 也均方可微，且

$$[aX(t)+bY(t)]' = aX'(t)+bY'(t)$$

(4) 设$\{X(t),t\in T\}$均方可微，$f(t)$为一个普通的可微函数，则$f(t)X(t)$也均方可微，且

$$\frac{\mathrm{d}}{\mathrm{d}t}[f(t)X(t)] = \frac{\mathrm{d}f(t)}{\mathrm{d}t}X(t)+f(t)\frac{\mathrm{d}X(t)}{\mathrm{d}t}$$

(5) 随机过程若均方可微，则必均方连续，反之不然.

显然，由均方可微的定义知

$$\begin{aligned}&\lim_{\Delta t\to 0}E\{[X(t+\Delta t)-X(t)]^2\}\\ =&\lim_{\Delta t\to 0}E\left\{\left[\frac{X(t+\Delta t)-X(t)}{\Delta t}\right]^2\right\}\cdot(\Delta t)^2 = E\{[X'(t)]^2\}\cdot 0 = 0\end{aligned}$$

(6) 若二阶矩过程$\{X(t),t\in T\}$均方可微，C为常数或随机变量，则

$$[X(t)+C]' = X'(t)$$

(7) 若二阶矩过程$\{X(t),t\in T\}$是n次均方可微的，则$\{X(t),t\in T\}$在t处的n阶均方导数的均值存在，且

$$E\left[\frac{\mathrm{d}^nX(t)}{\mathrm{d}t^n}\right] = \frac{\mathrm{d}^n}{\mathrm{d}t^n}E[X(t)]$$

即求均方导数与期望运算可交换顺序.

【例 7.8】 设有随机过程$\{X(t),t\in[0,1]\}$，定义

$$p\{X(0)=0\}=1,X(t)=Y_j,\frac{1}{2^j}<t\leqslant\frac{1}{2^{j-1}},j=1,2,\cdots$$

其中$\{Y_i\}$是独立同分布的随机变量序列，$E(Y_1)=0,D(Y_1)=1$，讨论$X(t)$的均方可微性.

解 由

$$E[X(t)]=0$$

$$\begin{aligned}R_X(s,t) &= E[X(s)X(t)]\\ &=\begin{cases}E(Y_j^{\,2})=1, & \dfrac{1}{2^j}<t,s\leqslant\dfrac{1}{2^{j-1}},j=1,2,\cdots\\ 0, & \text{其他}\end{cases}\end{aligned}$$

可知$\frac{\partial}{\partial s}R(0,0)=0=\frac{\partial}{\partial t}R(0,0),\frac{\partial^2}{\partial s\partial t}R(0,0)=\frac{\partial^2}{\partial t\partial s}R(0,0)=0$，但$R_X(s,t)$不是广义二次可微的，这是因为有

$$\begin{aligned}&\frac{\partial^2}{\partial s\partial t}R(s,t)=\frac{\partial^2}{\partial t\partial s}R(t,s)=E[X'(s)X'(t)]\\ =&\lim_{h,h_1\to 0}\frac{1}{h}\left[\lim_{h_1\to 0}\frac{R(s+h,t+h_1)-R(s+h,t)-R(s,t+h_1)+R(s,t)}{h_1}\right]\end{aligned}$$

若取$h_1=h$，则上式成为

$$\lim_{h\to 0}\frac{[R(h,h)-R(h,0)-R(0,h)+R(0,0)]}{h^2}=\lim_{h\to 0}\frac{1}{h^2}=+\infty$$

可知 $X(t)$ 不是均方可微的.

【例 7.9】 已知 $R_X(\tau)=\mathrm{e}^{-\tau^2}$，如果 $Y(t)=X(t)+X'(t)$，求 $R_Y(\tau)$.

解 依相关函数的定义，有

$$\begin{aligned}R_Y(\tau)&=E[Y(t)Y(t-\tau)]\\&=E\{[X(t)+X'(t)][X(t-\tau)+X'(t-\tau)]\}\\&=E\{X(t)X(t-\tau)+X'(t)X'(t-\tau)+X'(t)X(t-\tau)+X(t)X'(t-\tau)\}\\&=R_X(\tau)-R''_X(\tau)+R'_X(\tau)-R'_X(\tau)\\&=R_X(\tau)-R''_X(\tau)=(3-4\tau^2)\mathrm{e}^{-\tau^2}\end{aligned}$$

【例 7.10】 证明：维纳过程 $\{W(t),t\geqslant 0\}$ 是非均方可导的.

证明 $\forall t_0\in[0,+\infty)$，若维纳过程 $W(t)$ 在 t_0 处均方可导，则存在随机变量 Y，使得 $\lim\limits_{h\to 0}\dfrac{W(t_0+h)-W(t_0)}{h}=Y$，则 Y 均方可积，从而 $|Y|<+\infty$，且有数列 $\{h_0\}$，$0\leqslant h_n\to 0$，使得

$$\lim_{n\to+\infty}\frac{W(t_0+h_n)-W(t_0)}{h_n}=Y$$

所以

$$P\left\{\lim_{n\to+\infty}\frac{W(t_0+h_n)-W(t_0)}{\sqrt{h_n}}=0\right\}=1$$

故 $\forall\varepsilon>0$，$\exists N\geqslant 1$，当 $n\geqslant N$ 时，有 $P\left\{\left|\dfrac{X(t_0+h_n)-X(t_0)}{\sqrt{h_n}}\right|<\varepsilon\right\}>\dfrac{1}{2}$.

从另一方面知，因为 $W(t_0+h_n)-W(t_0)\sim N(0,\sigma^2 h_n)$，故

$$\frac{W(t_0+h_n)-W(t_0)}{\sqrt{h_n}}\sim N(0,\sigma^2)$$

由正态分布的性质知，当 ε 充分小时，式左边的概率为

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{|x|<\varepsilon}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}\mathrm{d}x<\frac{1}{2}$$

这与 h_n 无关，从而推出矛盾. 所以维纳过程是非均方可导的.

7.3 均方积分

本节主要介绍黎曼意义下的均方积分概念，进而简单介绍斯蒂尔吉斯均方积分的定义.

7.3.1 黎曼均方积分

定义7.8 设$\{X(t), t \in T\}$为二阶矩过程，$f(t)$为定义在T上的任意确定的函数，$[a, b] \subset T$. 将区间$[a, b]$分成n个子区间，分点为

$$a = t_0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n = b$$

记$\Delta_n = \max\limits_{1 \leqslant k \leqslant n}(t_{k-1} - t_k)$，作和式

$$S_n = \sum_{k=1}^{n} f(u_k) X(u_k)(t_k - t_{k-1}), \quad \forall u_k \in [t_{k-1}, t_k]$$

若$\underset{\Delta_n \to 0}{\mathrm{l.i.m}} S_n$存在且与$n$个子区间的分法及$u_k$的取法无关，则称此极限为$f(t)X(t)$在区间$[a, b]$上的均方积分，记为$\int_a^b f(t)X(t)\mathrm{d}t$. 此时也称$f(t)X(t)$在$[a, b]$上是(黎曼)均方积分可积的，即

$$\int_a^b f(t)X(t)\mathrm{d}t \triangleq \underset{\Delta_n \to 0}{\mathrm{l.i.m}} S_n = \underset{\Delta_n \to 0}{\mathrm{l.i.m}} \sum_{k=1}^{n} f(u_k) X(u_k)(t_k - t_{k-1})$$

特别地，当$f(t) \equiv 1, t \in [a, b]$，则

$$\int_a^b X(t)\mathrm{d}t = \underset{\Delta \to 0}{\mathrm{l.i.m}} \sum_{k=1}^{n} X(u_k)(t_k - t_{k-1})$$

称为随机过程$\{X(t), t \in T\}$在$[a, b]$上的均方积分.

定理7.8 (均方可积准则) 设随机过程$\{X(t), t \in T\}$，$f(t)$为定义在T上的任意确定的函数，$f(t)X(t)$在$[a, b]$上均方可积的充要条件为二重积分

$$\int_a^b \int_a^b f(s) f(t) R_X(s, t)\mathrm{d}s\mathrm{d}t$$

存在，其中$R_X(s, t)$是$\{X(t), t \in T\}$的自相关函数.

证明 充分性：若$f(s)f(t)R_X(s, t)$的二重积分存在，对$[a, b] \times [a, b]$的任意分割

$$a = s_0 < s_1 < \cdots < s_m = b, a = t_0 < t_1 < \cdots < t_n = b$$

及任意$(u_k, v_j) \in (s_{k-1}, s_k] \times (t_{j-1}, t_j]$ $(k = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$，有

$$\int_a^b \int_a^b f(s) f(t) R_X(s, t)\mathrm{d}s\mathrm{d}t = \lim_{\substack{\Delta s \to 0 \\ \Delta t \to 0}} \sum_{k=1}^{m} \sum_{j=1}^{n} f(u_k) f(v_j) R_X(u_k, v_j) \Delta s_k \Delta t_j$$

存在，其中$\Delta s = \max\limits_{1 \leqslant k \leqslant m} \Delta s_k$，$\Delta s_k = s_k - s_{k-1}$；$\Delta t = \max\limits_{1 \leqslant j \leqslant n} \Delta t_j$，$\Delta t_j = t_j - t_{j-1}$.

$$\text{上式} = \lim_{\substack{\Delta s\to 0\\ \Delta t\to 0}} \sum_{k=1}^{m}\sum_{j=1}^{n} E[f(u_k)X(u_k)f(v_j)X(v_j)]\Delta s_k \Delta t_j$$

$$= \lim_{\substack{\Delta s\to 0\\ \Delta t\to 0}} E\left[\sum_{k=1}^{m} f(u_k)X(u_k)\Delta s_k \sum_{j=1}^{n} f(v_j)X(v_j)\Delta t_j\right]$$

由均方收敛准则知

$$\underset{\Delta t\to 0}{\mathrm{l.i.m}} \sum_{k=1}^{n} f(u_k)X(u_k)\Delta t_k$$

存在，即 $f(t)X(t)$ 在$[a, b]$上均方可积.

必要性：由洛弗判别准则知，若均方积分 $\int_a^b f(t)X(t)\mathrm{d}t$ 存在，则下列极限存在，且

$$\lim_{\substack{\Delta s\to 0\\ \Delta t\to 0}} E\left[\sum_{k=1}^{m} f(u_k)X(u_k)\Delta s_k \sum_{j=1}^{n} f(v_j)X(v_j)\Delta t_j\right]$$

$$= E\left[\int_a^b f(s)X(s)\mathrm{d}s \int_a^b f(t)X(t)\mathrm{d}t\right]$$

$$= \int_a^b\int_a^b f(s)f(t)E[X(s)X(t)]\mathrm{d}s\mathrm{d}t = \int_a^b\int_a^b f(s)f(t)R_X(s,t)\mathrm{d}s\mathrm{d}t$$

推论 7.2　若随机过程$\{X(t), t\in T\}$的自相关函数$R_X(s,t)$在$[a,b]\times[a,b]$上可积，则 $X(t)$ 在$[a,b]$上均方可积，且

$$E\left[\left|\int_a^b X(t)\mathrm{d}t\right|^2\right] = \int_a^b\int_a^b R_X(s,t)\mathrm{d}s\mathrm{d}t$$

推论 7.3　若 $X(t)$ 在$[a,b]$上均方连续，则 $X(t)$ 在$[a,b]$上均方可积.

证明　根据均方连续性准则可知，$\{X(t), t\in T\}$在$[a,b]$上均方连续，$\{X(t), t\in T\}$的自相关函数$R_X(s,t)$在$[a,b]\times[a,b]$上连续，于是$R_X(s,t)$在$[a,b]\times[a,b]$上可积，进而有 $X(t)$ 在$[a,b]$上均方可积.

定义 7.9　（无穷区间上的广义均方积分）设$\{X(t), t\in T\}$为二阶矩过程，$f(t)$ 为定义在 T 上的任意确定的函数，考虑分别在以下三种情况下定义无穷区间上的均方积分：$T\in(-\infty, +\infty)$，$T\in(-\infty, b)$ 及 $T\in(a, +\infty)$.

(1) $\int_{-\infty}^{+\infty} f(t)X(t)\mathrm{d}t \triangleq \underset{a\to-\infty}{\mathrm{l.i.m}}\int_a^c f(t)X(t)\mathrm{d}t + \underset{b\to+\infty}{\mathrm{l.i.m}}\int_c^b f(t)X(t)\mathrm{d}t$；

(2) $\int_{-\infty}^{b} f(t)X(t)\mathrm{d}t \triangleq \underset{a\to-\infty}{\mathrm{l.i.m}}\int_a^b f(t)X(t)\mathrm{d}t$；

(3) $\int_{a}^{+\infty} f(t)X(t)\mathrm{d}t \triangleq \underset{b\to+\infty}{\mathrm{l.i.m}}\int_a^b f(t)X(t)\mathrm{d}t$.

类似地，可以得到无穷区间上的广义均方可积的充要条件：

定理 7.9 广义均方积分$\int_{-\infty}^{+\infty} f(t)X(t)\mathrm{d}t$存在的充分必要条件是广义二重积分

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(s)f(t)R_X(s,t)\mathrm{d}s\mathrm{d}t < +\infty$$

证明略.

定理 7.10 黎曼均方积分具有以下性质:

(1)(唯一性) 均方积分是唯一的,即如果

$$\int_a^b f(t)X(t)\mathrm{d}t = Y_1, \int_a^b f(t)X(t)\mathrm{d}t = Y_2$$

则有 $Y_1 = Y_2 (a.s)$.

(2)(线性性) 设$\{X(t), t \in T\}$、$\{Y(t), t \in T\}$是两个二阶矩过程,若它们在$[a, b] \subset T$上均方可积,A、B为常数或随机变量,则有

$$\int_a^b [Af(t)X(t) + Bg(t)Y(t)]\mathrm{d}t = A\int_a^b f(t)X(t)\mathrm{d}t + B\int_a^b g(t)Y(t)\mathrm{d}t$$

(3)(可加性) 若$\{X(t), t \in T\}$在$[a,b]$上均方可积,则

$$\int_a^b f(t)X(t)\mathrm{d}t = \int_a^c f(t)X(t)\mathrm{d}t + \int_c^b f(t)X(t)\mathrm{d}t, \quad a \leqslant c \leqslant b$$

(4)(不等性) 设$\{X(t), t \in T\}$在$[a, b]$上均方连续,则

$$\left\| \int_a^b X(t)\mathrm{d}t \right\| \leqslant \int_a^b \| X(t) \| \mathrm{d}t$$

证明
$$\begin{aligned}
E\left[\left|\int_a^b X(t)\mathrm{d}t\right|^2\right] &= \int_a^b\int_a^b R_X(s,t)\mathrm{d}s\mathrm{d}t \\
&\leqslant \int_a^b\int_a^b |R_X(s,t)|\mathrm{d}s\mathrm{d}t = \int_a^b\int_a^b |E[X(s)X(t)]|\mathrm{d}s\mathrm{d}t \\
&\leqslant \int_a^b\int_a^b \{E[|X(s)|^2]E[|X(t)|^2]\}^{\frac{1}{2}}\mathrm{d}s\mathrm{d}t \\
&= \left\{\int_a^b [E(|X(t)|^2)]^{\frac{1}{2}}\mathrm{d}t\right\}^2 = \left\{\int_a^b \|X(t)\|\mathrm{d}t\right\}^2
\end{aligned}$$

特别地,若$\{X(t), t \in T\}$在$[a, b]$上均方连续,则

$$E\left[\left|\int_a^b X(t)\mathrm{d}t\right|^2\right] \leqslant M(b-a)^2, M = \max_{a \leqslant x \leqslant b} E[|X(t)|^2]$$

(5)(直积性) 若$\{X(t), t \in T\}$在$[a, b]$及$[c, d]$上均方可积,则

$$E\left[\int_a^b X(t)\mathrm{d}t \cdot \int_c^d X(t)\mathrm{d}t\right] = \int_a^b\int_c^d R_X(s,t)\mathrm{d}s\mathrm{d}t$$

定义 7.10 若 $f(t)X(t)$ 在$[a,b]$上均方可积，令 $Y=\int_a^b f(t)X(t)\mathrm{d}t$，则均方积分的数学期望和二阶矩分别为

$$E(Y)=\int_a^b f(t)m_X(t)\mathrm{d}t;E(Y^2)=\int_a^b\int_a^b f(s)f(t)R_X(s,t)\mathrm{d}s\mathrm{d}t$$

事实上

$$\begin{aligned}E\left[\int_a^b f(t)X(t)\mathrm{d}t\right]&=E[\underset{\Delta\to0}{\mathrm{l.i.m}}\sum_{k=1}^n f(u_k)X(u_k)\Delta t_k]\\&=\lim_{\Delta\to0}E[\sum_{k=1}^n f(u_k)X(u_k)\Delta t_k]\\&=\lim_{\Delta\to0}\sum_{k=1}^n f(u_k)E[X(u_k)]\Delta t_k\\&=\lim_{\Delta\to0}\sum_{k=1}^n f(u_k)m_X(u_k)\Delta t_k=\int_a^b f(t)m_X(t)\mathrm{d}t\end{aligned}$$

定义 7.11 （黎曼均方变上限积分）设 $X(t)$ 在$[a,b]$上均方连续，对 $\forall t\in[a,b]$，有

$$Y(t)=\int_a^t X(s)\mathrm{d}s$$

称为 $X(t)$ 在$[a,b]$上的均方变上限积分.

类似一般函数的黎曼变上限积分的性质，不难得到黎曼均方变上限积分的性质如下：

设 $X(t)$ 在$[a,b]$上均方连续，则其在$[a,b]$上的均方变上限积分 $Y(t)$ 在$[a,b]$上均方可微，且有

(1) $\left[\int_a^t X(s)\mathrm{d}s\right]'=X(t)$；

(2) $E\left[\int_a^t X(s)\mathrm{d}s\right]=\int_a^t E[X(s)]\mathrm{d}s$（数学期望和积分可以交换顺序）；

(3) $R_Y(s,t)=\int_a^s\int_a^t R_X(u,v)\mathrm{d}u\mathrm{d}v$.

利用黎曼均方变上限积分的性质，类似于普通积分中的牛顿-莱布尼兹公式的推导，可以得到均方定积分的计算公式如下：

定理 7.11 （牛顿-莱布尼兹公式）设 $X(t)$ 在$[a,b]$上均方可导，$X'(t)$ 在$[a,b]$上均方连续，则有

$$\int_a^b X'(t)\mathrm{d}t=X(b)-X(a)$$

证明 因为 $X'(t)$ 在$[a,b]$上均方连续，所以有$\int_a^t X'(s)\mathrm{d}s$ 在$[a,b]$上均方可

导，且$\left[\int_a^t X'(s)\mathrm{d}s\right]' = X'(t)$，

即
$$\left[\int_a^t X'(s)\mathrm{d}s - X(t)\right]' = 0$$

故$\int_a^t X'(s)\mathrm{d}s - X(t) = C$，$C$为常数，$t \in [a,b]$，进而
$$\int_a^t X'(s)\mathrm{d}s = X(t) + C$$

令$t = a$，得$C = -X(a)$. 再令$t = b$得
$$\int_a^b X'(t)\mathrm{d}t = X(b) - X(a)$$

【例 7.11】 设$\{X(t), t \geqslant 0\}$是强度为λ的 Poisson 过程，令$M(T) = \frac{1}{T}\int_0^T X(t)\mathrm{d}t$，求$E[M(T)]$和$D[M(T)]$.

解 因为
$$\begin{aligned} R_X(s,t) &= E[X(s)X(t)] = C_X(s,t) - \lambda^2 st + \lambda^2 st + \lambda^2 st \\ &= \lambda\min\{s,t\} + \lambda^2 st \end{aligned}$$

所以$E[X^2(t)] = R_X(t,t) < +\infty$，且$R_X(s,t)$在$[0,T]\times[0,T]$上黎曼可积，并有
$$E[M(T)] = \frac{1}{T}\int_0^T E[X(t)]\mathrm{d}t = \frac{1}{T}\int_0^T \lambda t\,\mathrm{d}t = \frac{\lambda T}{2}$$
$$\begin{aligned} T^2 D[M(T)] &= \int_0^T\int_0^T C_X(s,t)\mathrm{d}s\mathrm{d}t = \int_0^T\int_0^T \lambda\min\{s,t\}\mathrm{d}s\mathrm{d}t \\ &= \lambda\left[\int_0^T\left(\int_0^t s\mathrm{d}s\right)\mathrm{d}t + \int_0^T\left(\int_t^T t\mathrm{d}s\right)\mathrm{d}t\right] = \frac{1}{3}\lambda T^3 \end{aligned}$$

故
$$D[M(T)] = \frac{1}{3}\lambda T$$

【例 7.12】 设$X(t) = 2A^2 t$，其中A是随机过程，$E(A^4) < +\infty$，求$\int_0^t X(t)\mathrm{d}t$.

解 因为$R_X(s,t) = E[X(s)X(t)] = E(4A^4 st) = 4stE(A^4)$连续，所以$\int_0^t X(t)\mathrm{d}t$存在. 若在定义中取$u_k = \frac{1}{2}(t_{k+1} + t_k)$，则
$$\begin{aligned} \sum_{k=0}^{n-1} X(u_k)(t_{k+1} - t_k) &= \sum_{k=0}^{n-1} 2A^2 \frac{1}{2}(t_{k+1} + t_k)(t_{k+1} - t_k) \\ &= A^2(t^2 - 0) = A^2 t^2 \end{aligned}$$

故
$$\int_0^t X(t)\mathrm{d}t = \int_0^t 2A^2 t\mathrm{d}t = A^2 t^2$$

7.3.2　斯蒂尔吉斯(Stieltjes)均方积分

定义 7.12　(Stieltjes 均方积分) 设 $\{X(t), t \in T\}$ 为二阶矩过程，$f(t)$ 为定义在 T 上的普通函数，$T = [a,b]$. 将区间 $[a,b]$ 进行分割

$$a = t_0 < t_1 < \cdots < t_n = b$$

记 $\Delta_n = \max\limits_{1 \leqslant k \leqslant n}(t_{k-1} - t_k)$，作和式

$$Z_n = \sum_{k=1}^{n} f(u_k)[X(t_k) - X(t_{k-1})], \text{其中 } u_k \in [t_{k-1}, t_k], k = 1,2,\cdots,n$$

如果 $\Delta_n \to 0$ 时，$\underset{\Delta_n \to 0}{\text{l.i.m}}(Z_n) = Z$ 存在，则称 Z 为 $f(t)$ 关于 $X(t)$ 在 $[a,b]$ 上的**斯蒂尔吉斯均方积分**，记为 $Z = \int_a^b f(t)\mathrm{d}X(t)$.

以下不加证明地给出斯蒂尔吉斯均方积分的可积性准则和积分性质：

定理 7.12　(可积性准则) 设 $\{X(t), t \in [a,b]\}$ 为二阶矩过程，其相关函数为 $R_X(t,s)$，$f(t)$ 为定义在 $[a,b]$ 上的普通函数. 若二重斯蒂尔吉斯均方积分

$$\int_a^b\int_a^b f(s)f(t)\mathrm{d}R_X(s,t) < +\infty$$

则 $f(t)$ 关于 $X(t)$ 在 $[a,b]$ 上的斯蒂尔吉斯均方积分 $\int_a^b f(t)\mathrm{d}X(t)$ 存在.

定理 7.13　若 $f(t)$ 关于 $X(t)$ 在 $[a,b]$ 上的斯蒂尔吉斯均方可积，则有

(1) $E\left[\int_a^b f(t)\mathrm{d}X(t)\right] = \int_a^b f(t)\mathrm{d}E[X(t)] = \int_a^b f(t)\mathrm{d}m_X(t)$;

(2) $E\left[\left|\int_a^b f(t)\mathrm{d}X(t)\right|^2\right] = \int_a^b\int_a^b f(s)\,\overline{f(t)}\mathrm{d}R_X(s,t)$.

有限区间上的斯蒂尔吉斯均方积分的定义和结果也可以类似地推广到无限区间的情况，读者可以自行完成，此处略.

习　题　7

1. 设 $\{X(t), t \in [a,b]\}$ 为二阶矩过程，均值为 $E[X(t)] = \alpha + \beta t$，协方差函数为 $C_X(t_1,t_2) = \mathrm{e}^{-\lambda|t_1-t_2|}$. 令 $Y(t) = X(t+1) - X(t)$，试证 $Y(t)$ 为平稳过程，并求它的均值和相关函数.
2. 设有二阶随机序列 $\{X(n)\}$，$\{X(n)\}$ 的相关函数为 $R_X(n_1,n_2) = E[X(n_1)\overline{X(n_2)}]$. 如果有序列 $\{a_n\}$，$n = 1,2,\cdots$，并定义 $Y(n) = \sum\limits_{k=1}^{n} a_k X(k)$，则在什么条

件下，$Y(n)$ 是均方收敛序列？

3. 设 $\{X(t)\}$ 是平稳过程，相关函数为 $R_X(\tau)$. 令 $Y=\int_a^{a+T}X(t)\mathrm{d}t$，其中 $T>0$，a 是实数. 试证 $E(|Y|^2)=\int_{-T}^{T}(T-|\tau|)R_X(\tau)\mathrm{d}\tau$.
4. 设有二阶随机序列 $\{X(n)\}$ 和二阶矩变量 X，$f(n)$ 是普通函数，且满足李普希兹条件 $|f(u)-f(v)|\leqslant M|u-v|$，其中 M 是一正常数，证明：
$$\mathrm{l.i.m}\, f(X_n)=f(X)$$
5. 证明：$\{X(t),t\in T\}$ 在 $t=0$ 处均方连续与其协方差函数 $C_X(\tau)$ 在 $\tau=0$ 处连续等价.
6. 设 $\{X(t),t\in T\}$ 是实均方随机过程，求其导数过程 $\{X'(t),t\in T\}$ 的协方差函数 $C_{X'}(s,t)$.

参考答案

习题 2

1. $X(t)$ 的密度函数

$$f_X(t)=\frac{2x}{t^2}\mathrm{e}^{-\left(\frac{x}{t}\right)^2}$$

所以均值函数

$$m_X(t)=\int_0^{+\infty}\frac{2x^2}{t^2}\mathrm{e}^{-\left(\frac{x}{t}\right)^2}\mathrm{d}x=2t\int_0^{+\infty}\left(\frac{x}{t}\right)^2\mathrm{e}^{-\left(\frac{x}{t}\right)^2}\mathrm{d}\,\frac{x}{t}=2t\int_0^{+\infty}\mu^2\mathrm{e}^{-\mu^2}\mathrm{d}\mu$$

$$=t\int_0^{+\infty}u^{\frac{1}{2}}\mathrm{e}^{-u}\mathrm{d}u=t\Gamma\left(\frac{3}{2}\right)$$

与 t 有关，所以$\{X(t),t>0\}$ 不是弱平稳过程.

2. $X(t)$ 的密度函数

$$f_t(x)=\frac{1}{\sqrt{2\pi t}\sigma}\mathrm{e}^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}$$

所以均值函数

$$m_X(t)=\int_{-\infty}^{+\infty}x\,\frac{1}{\sqrt{2\pi t}\sigma}\mathrm{e}^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}\mathrm{d}x=\mu t$$

方差函数

$$\mathrm{Var}[X(t)]=\int_{-\infty}^{+\infty}x^2\,\frac{1}{\sqrt{2\pi t}\sigma}\mathrm{e}^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}\mathrm{d}x=\sigma^2 t.$$

3. (1) 均值函数

$$m_X(t)=E[X(t)]=E(A)E[\cos(\omega_0 t+\Phi)]$$

$$=E(A)\,\frac{1}{2\pi}\int_0^{2\pi}\sin(\omega t+\varphi)\mathrm{d}\varphi=-E(A)\,\frac{1}{2\pi}[\cos(\omega t+\varphi)]\Big|_0^{2\pi}=0$$

相关函数

$$R(s,t)=E\{[A\sin(\omega s+\Phi)][A\sin(\omega t+\Phi)]\}$$

$$=E(A^2)\,\frac{1}{2\pi}\int_0^{2\pi}\sin(\omega s+\varphi)\sin(\omega t+\varphi)\mathrm{d}\varphi$$

$$=E(A^2)\,\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}\{\cos[\omega(t-s)]-\cos[\omega(s+t)+2\varphi]\}\mathrm{d}\varphi$$

$$=\frac{1}{2}E(A^2)\cos[\omega(t-s)]$$

协方差函数

$$C(s,t)=R(s,t)-0=\frac{1}{2}E(A^2)\cos[\omega(t-s)]$$

(2) 当 $t=0$ 和 $t=1$ 时，$X(0)=A\sin\Phi$ 与 $X(1)=A\sin(\omega+\Phi)$ 不同分布，故$\{X(t),t\in(-\infty,$

$+\infty)\}$ 不是强平稳的. 但由(1)知 $m_X(t)=0$; $E[X^2(t)]=R(t,t)=\frac{1}{2}E(A^2)<+\infty$;

$C(s,t)=\frac{1}{2}E(A^2)\cos[\omega(t-s)]$ 仅与时间差 $t-s$ 有关,所以 $\{X(t),t\in(-\infty,+\infty)\}$ 是弱平稳的.

4. 由 2.3 知,若 $\{A(t),t\in(-\infty,+\infty)\}$ 是弱平稳过程,则

(1) $X(t)$ 的均值函数 $m_X(t)=0$.

(2) $X(t)$ 的相关函数

$$\begin{aligned}R(s,t)&=E\{[A(s)\sin(\omega s+\Phi)][A(t)\sin(\omega t+\Phi)]\}\\&=E[A(s)A(t)]\frac{1}{2\pi}\int_0^{2\pi}\sin(\omega s+\varphi)\sin(\omega t+\varphi)\mathrm{d}\varphi\\&=E[A(s)A(t)]\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}\{\cos[\omega(t-s)]-\cos[\omega(s+t)+2\varphi]\}\mathrm{d}\varphi\\&=\frac{1}{2}\cos[\omega(t-s)]E[A(s)A(t)]\end{aligned}$$

当 $\{A(t),t\in(-\infty,+\infty)\}$ 是弱平稳过程时, $E[A(s)A(t)]$ 只与 $t-s$ 有关,故 $R(s,t)$ 只与 $t-s$ 有关.

(3) $E[X^2(t)]=R(t,t)=\frac{1}{2}E[A^2(t)]<+\infty$.

由上知, $\{X(t),t\in(-\infty,+\infty)\}$ 是弱平稳过程.

5.

$X(t)$ 的均值函数

$$m_X(t)=\sum_{i=1}^n a_iE[\sin(\omega t+\Phi_i)]=0$$

$X(t)$ 的相关函数

$$\begin{aligned}R(s,t)&=E[X(s)X(t)]=E\{[\sum_{i=1}^n a_i\sin(\omega t+\Phi_i)][\sum_{i=1}^n a_i\sin(\omega s+\Phi_i)]\}\\&=E[\sum_{i=1}^n a_i^2\sin(\omega t+\Phi_i)\sin(\omega s+\Phi_i)+\sum_{i\neq j}a_ia_j\sin(\omega t+\Phi_i)\sin(\omega s+\Phi_j)]\\&=\frac{1}{2}\sum_{i=1}^n a_i^2\cos[\omega(t-s)]\end{aligned}$$

$X(t)$ 的协方差函数

$$C(s,t)=R(s,t)-m_X(t)m_X(s)=\frac{1}{2}\sum_{i=1}^n a_i^2\cos[\omega(t-s)]$$

6. (2) $C(s,t)=\begin{cases}1, & n\leqslant s,t\leqslant n+\frac{1}{2},n=0,1,2,\cdots\\0, & \text{其他}\end{cases}$

(3) 不是

7. 证明:

(1) $m_Z(t)=\cos(\omega t)E[X(t)]-\sin(\omega t)E[Y(t)]=0$

(2) $Z(t)$ 的相关函数

$$R_Z(s,t)=E[Z(t)Z(s)]=\cos(\omega t)\cos(\omega s)E[X(t)X(s)]+\sin(\omega t)\sin(\omega s)E[Y(t)Y(s)]$$
$$-\cos(\omega t)\sin(\omega s)E[X(t)Y(s)]-\cos(\omega s)\sin(\omega t)E[X(s)Y(t)]$$
$$=\cos(\omega t)\cos(\omega s)C(t-s)+\sin(\omega t)\sin(\omega s)C(t-s)$$
$$=\cos[\omega(t-s)]C(t-s)$$

(3)$Z(t)$ 的协方差函数

$$C_Z(s,t)=R_Z(s,t)-m_Z(t)m_Z(s)=\cos[\omega(t-s)]C(t-s)$$

8.(1) ①$E[X^2(t)]=E[\sin^2(\Phi t)]\leqslant 1$

②$m_X(t)=E[\sin(\Phi t)]=0$

③$C(s,t)=E[\sin(\Phi t)\sin(\Phi s)]$

$$=\int_0^{2\pi}\frac{1}{2\pi}\sin(ut)\sin(us)\mathrm{d}u$$
$$=\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}\{\cos[u(t-s)]-\cos[u(t+s)]\}\mathrm{d}u$$
$$=\begin{cases}\frac{1}{2}, & t=s\\ 0, & t\neq s\end{cases},t,s\in\mathbf{Z}$$

由 ①,②,③ 知$\{X(t),t=1,2,\cdots\}$是弱平稳过程.

当 $t=1$ 和 $t=2$ 时,$X(1)=\sin\Phi$ 与 $X(2)=\sin(2\Phi)$ 不同分布,故$\{X(t),t=1,2,\cdots\}$不是强平稳的.

(2) $$E[X(t)]=\begin{cases}0, & t=0\\ \frac{1-\cos(2\pi t)}{2\pi t}, & t\neq 0\end{cases},$$

故$\{X(t),t\geqslant 0\}$不是宽平稳过程,而 $X(0)$ 与 $X(1)=\sin\Phi$ 的分布函数不同,故$\{X(t),t\geqslant 0\}$也不是严平稳过程.

习题 3

1. $$P\left\{X\left(\frac{1}{2}\right)\geqslant 2\right\}=\sum_{i=2}^{+\infty}\frac{\left(3\cdot\frac{1}{2}\right)^i}{i!}\mathrm{e}^{-\frac{3}{2}}=1-\frac{\left(\frac{3}{2}\right)^0}{0!}\mathrm{e}^{-\frac{3}{2}}-\frac{\left(\frac{3}{2}\right)^1}{1!}\mathrm{e}^{-\frac{3}{2}}=0.4422.$$

2.(1) 假设 $s>t$

$$E[N(t)N(s)]=E\{N(t)[N(s)-N(t)+N(t)]\}$$
$$=E\{N(t)[N(s)-N(t)]\}+E[N^2(t)]$$
$$=E[N(t)]E[N(s)-N(t)]+D[N(t)]+\{E[N(t)]\}^2$$
$$=\lambda^2t(s-t)+\lambda t+\lambda^2t^2=\lambda^2ts+\lambda t$$

所以,协方差函数

$$C(s,t)=E[N(t)N(s)]-E[N(t)]E[N(s)]=\lambda^2ts+\lambda t-\lambda t\cdot\lambda s=\lambda t$$

同理可证当 $s<t$ 时,$C(s,t)=\lambda s$,

故 $C(s,t)=\lambda\min(s,t)$.

3. 由齐次 Poisson 过程的可分解性知,12:00 ~ 14:00 通过该十字路口,但是忽视指示牌“停”的车辆,服从强度为 $\lambda=40\times0.8\%=0.32$(辆/h) 的齐次 Poisson 过程,其以 12:00 作为 0 时刻,则在 12:00 ~ 13:00 间至少有一辆汽车忽视指示牌的概率为

$$P\{N(1)-N(0)=0\}=1-\frac{(0.32)^0}{0!}e^{-0.32}=0.2739$$

5.(1) 在 12:00 ~ 16:00 之间到达加油站的平均车辆为

$$m(4)=\int_0^4\lambda(t)\mathrm{d}t=\int_0^4(8-4t+3t^2)\mathrm{d}t=64$$

(2) 在 12:00 ~ 16:00 之间至少有 40 辆车达到的概率为

$$P\{N(4)-N(1)<40\}=1-\sum_{i=1}^{39}P\{N(4)-N(1)=i\}=1-\sum_{i=1}^{39}\frac{64^i}{i!}e^{-i}=0.89$$

6.(1)
$$\begin{aligned}E\{N(t)N(s+t)\}&=E\{N(t)[N(s+t)-N(t)+N(t)]\}\\&=E\{N(t)[N(s+t)-N(t)]\}+E[N^2(t)]\\&=E[N(t)]E[N(s)]+D[N(t)]+\{E[N(t)]\}^2\\&=\lambda^2ts+\lambda t+\lambda^2t^2\end{aligned}$$

(2)
$$\begin{aligned}E[N(t+s)\mid N(s)=m]&=E[N(t+s)-N(s)+N(s)\mid N(s)=m]\\&=E[N(t+s)-N(s)\mid N(s)=m]+E[N(s)\mid N(s)=m]\\&=\lambda t+m\end{aligned}$$

因为 $P\{E[N(t+s)\mid N(s)]=E[N(t+s)\mid N(s)=m]=P\{N(s)=m\}$

所以 $E[N(t+s)\mid N(s)]$ 的分布列为

$$P\{E[N(t+s)\mid N(s)]=\lambda t+m\}=\frac{(\lambda s)^m}{m!}e^{-\lambda s}$$

(3) 令 $t=s+\Delta,\Delta\geqslant 0$

$$P\{N(s)\leqslant N(t)\}=P\{N(s+\Delta)-N(s)\geqslant 0\}=\sum_{k=0}^{+\infty}\frac{(\lambda\Delta)^k}{k!}e^{-\lambda\Delta}=1$$

(4)$0\leqslant P\{N(t)-N(s)>\varepsilon\}\leqslant P\{N(t)-N(s)\geqslant 1\}$

而

$$\begin{aligned}\lim_{t\to s}P\{N(t)-N(s)\geqslant 1\}&=1-\lim_{t\to s}P\{N(t)-N(s)=0\}\\&=1-\lim_{t\to s}e^{-\lambda(t-s)}=0\end{aligned}$$

所以$\lim\limits_{t\to s}P\{N(t)-N(s)>\varepsilon\}=0$.

7.解 由 $T\sim e(\lambda)$ 知

$$\begin{aligned}P\{T>t_1+t_2\mid T>t_1\}&=\frac{P\{T>t_1+t_2,T>t_1\}}{P\{T>t_1\}}\\&=\frac{e^{-\lambda(t_1+t_2)}}{e^{-\lambda t_1}}=e^{-\lambda t_2}=P\{T>t_2\}\end{aligned}$$

8.$N_1(t)$ 的到达时间间隔 W_n 的密度函数为

$$f_w(t)=\begin{cases}\lambda_1e^{-\lambda_1t}, & t\geqslant 0\\0, & \text{其他}\end{cases}$$

$$P\{N_2(t)=k\}=\frac{(\lambda_2t)^k}{k!}e^{-\lambda_2t}$$

所以

$$\begin{aligned}p&=P\{N_2(T_n)-N_2(T_{n-1})=k\}=P\{N_2(T_n-T_{n-1})=k\}=P\{N_2(W_n)=k\}\\&=\int_0^{+\infty}P\{N_2(t)=k\}f_w(t)\mathrm{d}t=\int_0^{+\infty}\frac{(\lambda_2t)^k}{k!}e^{-\lambda_2t}\lambda_1e^{-\lambda_1t}\mathrm{d}t\\&=\frac{\lambda_1\lambda_2^k}{k!}\int_0^{+\infty}t^ke^{-(\lambda_1+\lambda_2)t}\mathrm{d}t=\frac{\lambda_1\lambda_2^k}{k!}\cdot\frac{\Gamma(k+1)}{(\lambda_1+\lambda_2)^{k+1}}=\frac{\lambda_1}{\lambda_1+\lambda_2}\cdot(\frac{\lambda_2}{\lambda_1+\lambda_2})k\end{aligned}$$

习题 4

1. $\boldsymbol{P}=\begin{pmatrix} p & q & 0 & 0 \\ 0 & 0 & p & q \\ p & q & 0 & 0 \\ 0 & 0 & q & p \end{pmatrix}$, $\boldsymbol{P}^{(2)}=\begin{pmatrix} p^2 & pq & pq & q^2 \\ p^2 & pq & pq & q^2 \\ p^2 & pq & pq & q^2 \\ p^2 & pq & pq & q^2 \end{pmatrix}$

4. $\boldsymbol{P}^{(3)}=\begin{pmatrix} 0.25 & 0.375 & 0.375 \\ 0.375 & 0.25 & 0.375 \\ 0.375 & 0.375 & 0.25 \end{pmatrix}$, $p_2(3)=0.25$

5. $\boldsymbol{P}^T=(0.42,0.26,0.32)$, $\boldsymbol{P}^T=(0.426,0.288,0.286)$

6. (1) $C_1=\{1,2,3\}$, $C_2=\{4,5\}$ 为两个遍历状态闭集.

(2) $C=\{1,2,3\}$ 为常返闭集, $N=\{4\}$ 为非常返态.

(3) $C_1=\{0\}$, $C_2=\{b\}$ 是吸收态闭集, $N=\{1,2,\cdots,b-1\}$ 是非常返集.

7. $N=\{1,2\}$ 是非常返集, $C_1=\{3,4,5\}$, $C_2=\{6,7\}$ 是正常返闭集. 由转移矩阵

$$\begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.4 & 0 & 0.6 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

可解得 C_1 的平稳分布为 $\{0,0,\frac{10}{23},\frac{7}{23},\frac{6}{23},0,0\}$.

同理可得 C_2 的平稳分布为 $\{0,0,0,0,0,\frac{8}{15},\frac{7}{15}\}$.

8. (1) $f_{11}^{(1)}=\frac{1}{2}$, $f_{11}^{(2)}=\frac{1}{6}$, $f_{11}^{(3)}=\frac{1}{9}$, $f_{12}^{(1)}=\frac{1}{2}$, $f_{12}^{(2)}=\frac{1}{4}$, $f_{12}^{(3)}=\frac{1}{8}$

(2) $f_{11}^{(1)}=p_1$, $f_{11}^{(2)}=0$, $f_{11}^{(3)}=q_1q_2q_3$, $f_{12}^{(1)}=q_1$, $f_{12}^{(2)}=p_1q_1$, $f_{12}^{(3)}=p_1^2q_1$.

9. $\pi_j=\frac{p_1\cdots p_{j-1}}{q_1\cdots q_j}\pi_0$, $j\geqslant 1$, $\pi_0=\frac{1}{1+\sum\limits_{j=1}^{+\infty}\prod\limits_{k=0}^{j-1}\frac{p_k}{q_{k+1}}}$.

10. (1) $\boldsymbol{P}=\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{2}{9} & \frac{5}{9} & \frac{2}{9} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$

(2) 由于 $I=\{0,1,2\}$ 是有限链, I 中所有状态是互通的, 且状态 0 是非周期的, 故 $\{X_n\}$ 为遍历链.

(3) $\pi_0=\frac{1}{5}$, $\pi_1=\frac{3}{5}$, $\pi_2=\frac{1}{5}$.

$\lim\limits_{n\to+\infty}p_{i_0}^{(n)}=\pi_0=\frac{1}{5}$, $\lim\limits_{n\to+\infty}p_{i_1}^{(n)}=\pi_1=\frac{3}{5}$, $\lim\limits_{n\to+\infty}p_{i_2}^{(n)}=\pi_2=\frac{1}{5}$.

13. (2) $\pi_1=0.1788$, $\pi_2=0.3613$, $\pi_3=0.3426$, $\pi_4=0.11173$.

(3) $\mu_4=8.53$(天).

习题 5

3. 由题设知，柯尔莫哥洛夫向前方程为

$$p'_{ij}(t)=-p_{ij}(t)+0.5p_{i,j-1}(t)+0.5p_{i,j+1}(t)$$

由于状态空间为 $I=\{1,2,3\}$，故

$$p_{ij}(t)+p_{i,j-1}(t)+p_{i,j+1}(t)=1$$

所以

$$p'_{ij}(t)=-p_{ij}(t)+0.5[1-p_{ij}(t)]$$

$$=-\frac{3}{2}p_{ij}(t)+0.5$$

解上述微分方程得

$$p_{ij}(t)=ce^{-\frac{3}{2}t}+\frac{1}{3}$$

由初始条件

$$p_{ij}(0)=\begin{cases}1, & i=j\\0, & i\neq j\end{cases}$$

确定常数 c 得

$$p_{ij}(t)=\begin{cases}\frac{1}{3}-\frac{1}{3}e^{-\frac{3}{2}t}, & i=j\\[2ex]\frac{1}{3}+\frac{1}{3}e^{-\frac{3}{2}t}, & i\neq j\end{cases}$$

故其平稳分布为

$$\pi_j=\lim_{t\to+\infty}p_{ij}(t)=\frac{1}{3},\quad j=1,2,3$$

4. 设服务员的服务时间为 T，则由题意知 T 服从指数分布，其概率密度函数为

$$f(t)=\begin{cases}\mu e^{-\mu t}, & t>0\\0, & t\leqslant 0\end{cases}$$

记$[0,t]$内到达的顾客数为 $X(t)$，则

$$P\{X(t)=n\}=\frac{(\lambda t)^n}{n!}e^{-\lambda t},n=0,1,2,\cdots$$

(1) 在服务员的服务时间内到达的顾客的平均数为

$$E\{E[X(t)]\mid T=t\}=\int_0^{+\infty}\left[\sum_{n=0}^{+\infty}n\frac{(\lambda t)^n}{n!}e^{-\lambda t}\right]\mu e^{-\mu t}dt$$

$$=\int_0^{+\infty}\lambda t\mu e^{-\mu t}dt=\frac{\lambda}{\mu}$$

(2) 在服务员的服务时间内无顾客到达的概率为

$$p_0=\int_0^{+\infty}e^{-\lambda t}\cdot\mu e^{-\mu t}dt=\frac{\mu}{\mu+\lambda}$$

5. (1) 记机器正常工作状态为 0，维修状态为 1，则状态空间为$\{0,1\}$.

(2) 由指数分布的“无记忆性”，可求得 Q 矩阵为

$$Q=\begin{pmatrix}-\lambda & \lambda\\ \mu & -\mu\end{pmatrix}$$

(3) 向前或向后退方程及初始条件为

$$\begin{pmatrix} p'_{00}(t) \\ p'_{10}(t) \end{pmatrix} = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \begin{pmatrix} p_{00}(t) \\ p_{10}(t) \end{pmatrix}$$

$$\begin{pmatrix} p_{00}(0) \\ p_{10}(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$P\{X(0)=0\} = p_0(0) = 1$$

$$P\{X(0)=1\} = p_1(0) = 0$$

(4) 解上面的微分方程

$$p_{00}(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}$$

$$p_{10}(t) = \frac{\mu}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}$$

由

$$P\{X(t)=j\} = p_j(t) = \sum_{i\in I} p_i(0)p_{ij}(t) = p_{0j}(t)$$

可求得

$$p_0(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}$$

$$p_0(10) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-10(\lambda+\mu)}$$

(5) 根据以上所求,求极限即有

$$\lim_{t\to+\infty} p_{00}(t) = \frac{\mu}{\lambda+\mu} \triangleq p_0$$

$$\lim_{t\to+\infty} p_{10}(t) = \frac{\mu}{\lambda+\mu} \triangleq p_0$$

$$\lim_{t\to+\infty} p_0(t) = \frac{\mu}{\lambda+\mu} \triangleq p_0$$

另外根据上面的极限性态讨论,由

$$\begin{cases} \sum_{k\in I} p_k q_{kj} = 0 \\ \sum_{k\in I} p_k = 1 \end{cases}$$

同样可以求得

$$\lim_{t\to+\infty} p_0(t) = p_0 = \frac{\mu}{\lambda+\mu}$$

$$\lim_{t\to+\infty} p_1(t) = p_1 = \frac{\lambda}{\lambda+\mu}$$

6. 状态空间为 $S=\{0,1,2,\cdots,n\}$, $k=0,1,2,\cdots,n$,并且是一生灭过程,生灭矩阵为

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots & 0 \\ \mu & -(\lambda+\mu) & \lambda & 0 & \cdots & 0 \\ 0 & 2\mu & -(\lambda+2\mu) & \lambda & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & (n-1)\mu & -[\lambda+(n-1)\mu] & \lambda \\ 0 & 0 & 0 & 0 & n\mu & -n\mu \end{pmatrix}$$

由 Q 矩阵得绝对概率分布方程为

$$
\begin{cases}
\dfrac{\mathrm{d}p_0(t)}{\mathrm{d}t} = -\lambda p_0(t) + \mu p_1(t) \\
\dfrac{\mathrm{d}p_1(t)}{\mathrm{d}t} = \lambda p_0(t) - (\lambda + \mu) p_1(t) + 2\mu p_2(t) \\
\quad \cdots \\
\dfrac{\mathrm{d}p_k(t)}{\mathrm{d}t} = \lambda p_{k-1}(t) - (\lambda + k\mu) p_k(t) + (k+1)\mu p_{k+1}(t) \quad (0 < k < n) \\
\quad \cdots \\
\dfrac{\mathrm{d}p_n(t)}{\mathrm{d}t} = \lambda p_{n-1}(t) - n\mu p_n(t)
\end{cases}
$$

令 $t \to +\infty$,有

$$
\begin{cases}
-\lambda p_0 + \mu p_1 = 0 \\
\quad \cdots \\
\lambda p_{k-1} - (\lambda + k\mu) p_k + (k+1)\mu p_{k+1} = 0 \quad (0 < k < n) \\
\quad \cdots \\
\lambda p_{n-1} - n\mu p_n = 0
\end{cases}
$$

设 $g_k = \lambda p_{k-1} - k\mu p_k$,则

$$
g_k = \lambda p_{k-1} - k\mu p_k = \lambda p_k - (k+1)\mu p_{k+1} = g_{k+1}
$$
$$
g_0 = 0, g_n = 0
$$

因此 $g_i = 0, i = 0,1,2,\cdots,n$

故

$$
\begin{cases}
p_1 = \dfrac{\lambda}{\mu} p_0 \\
\quad \cdots \\
p_k = \dfrac{\lambda}{k\mu} p_{k-1} = \dfrac{1}{k!} \left(\dfrac{\lambda}{\mu}\right)^k p_0 \\
\quad \cdots \\
p_n = \dfrac{1}{n!} \left(\dfrac{\lambda}{\mu}\right)^n p_0
\end{cases}
$$

利用 $\sum\limits_{k=0}^{n} p_k = 1 \Rightarrow p_0 \left[1 + \dfrac{\lambda}{\mu} + \dfrac{1}{2!}\left(\dfrac{\lambda}{\mu}\right)^2 + \cdots + \dfrac{1}{n!}\left(\dfrac{\lambda}{\mu}\right)^n\right] = 1$

可得

$$
p_k = \frac{\dfrac{1}{k!}\left(\dfrac{\lambda}{\mu}\right)^k}{\sum\limits_{i=0}^{n} \dfrac{1}{i!}\left(\dfrac{\lambda}{\mu}\right)^i}, 0 \leqslant k \leqslant n.
$$

习题 6

5. $f_t(x \mid B(s) = y) = \dfrac{1}{\sqrt{2\pi(t-s)}\sigma} \exp\{-\dfrac{(x-y)^2}{2(t-s)\sigma^2}\}, \quad 0 \leqslant s < t$

6. 证明　因为 $\{B_1(t), t \geqslant 0\}$ 与 $\{B_2(t), t \geqslant 0\}$ 为相互独立的标准 Brown 运动,所以 $X(0) = B_1(0) - B_2(0) = 0$,且 $X(t) - X(s) = [B_1(t) - B_1(s)] - [B_2(t) - B_2(s)]$.

由 $B_1(t)$、$B_2(t)$ 的独立平稳增量性可知，$\{X(t),t\geq 0\}$ 也具有独立平稳增量.

又 $E[X(t)]=E[B_1(t)]-E[B_2(t)]=0$，$\mathrm{Var}[X(t)]=\mathrm{Var}[B_1(t)]+\mathrm{Var}[B_2(t)]=2t$，所以 $\{X(t)\}$ 是布朗运动.

7. 证明　对于 $s\leqslant t$，有

$$\mathrm{Cov}[X(s),X(t)]=(1+s)(1+t)\mathrm{Cov}[Z\left(\frac{t}{1+t}\right),Z\left(\frac{s}{1+s}\right)],$$

由于 $\{Z(t)\}$ 与 $\{B(t)-B(1)\}$ 同分布，其中 $\{B(t)\}$ 是 Brown 运动，可见

$$\begin{aligned}\mathrm{Cov}[X(s),X(t)]&=(1+s)(1+t)\{\mathrm{Cov}[B\left(\frac{t}{1+t}\right),B\left(\frac{t}{1+t}\right)]-\frac{t}{1+t}\mathrm{Cov}[B(1),B\left(\frac{s}{1+s}\right)]\\&-\frac{s}{1+s}\mathrm{Cov}[B\left(\frac{t}{1+t}\right),B(1)]+\frac{st}{(1+s)(1+t)}\mathrm{Cov}[B(1),B(1)]\}\\&=s(1+t)-st-st+st=s(s\leqslant t)\end{aligned}$$

易见 $\{X(t)\}$ 是高斯过程，$E[X(t)]=0$，故 $\{X(t)\}$ 是布朗运动.

8. 考虑随机向量 $\boldsymbol{X}=(B(1),\cdots,B(n))'$，由于 $\{B(t)\}$ 是标准 Brown 运动，所以 X 服从多元正态分布，具有零均值，其协方差矩阵为

$$\sum=\begin{pmatrix}1&1&1&1&\cdots&1&1&1\\1&2&2&2&\cdots&2&2&2\\1&2&3&3&\cdots&3&3&3\\\vdots&\vdots&\vdots&\vdots&&\vdots&\vdots&\vdots\\1&2&3&4&\cdots&n-2&n-1&n-1\\1&2&3&4&\cdots&n-2&n-1&n\end{pmatrix}_{n\times n}$$

令 $\boldsymbol{A}=(1,1,\cdots,1)$，则 $\boldsymbol{AX}=B(1)+B(2)+\cdots+B(n)$ 是具有零均值的随机变量，方差为 $\boldsymbol{A}\sum\boldsymbol{A}'=\dfrac{n(n+1)(2n+1)}{6}$. 又因为 $X(t)=tB\left(\dfrac{1}{t}\right)$，所以 $X(0)=0$，于是

$$\begin{aligned}X(t)-X(s)&=tB\left(\frac{1}{t}\right)-sB\left(\frac{1}{s}\right)\\&=tB\left(\frac{1}{t}\right)-tB\left(\frac{1}{s}\right)+tB\left(\frac{1}{s}\right)-sB\left(\frac{1}{s}\right)\\&=t[B\left(\frac{1}{t}\right)-B\left(\frac{1}{s}\right)]+(t-s)B\left(\frac{1}{s}\right)\sim N(0,|t-s|)\end{aligned}$$

所以 $X(t)$ 是独立平稳增量过程，又 $X(t)=tB\left(\dfrac{1}{t}\right)\sim N(0,t)$，故 $\{X(t),t\geqslant 0\}$ 是布朗运动.

9. 因为

$$E[X(t)]=(1-t)E[B\left(\frac{t}{1-t}\right)]=0$$

$$\begin{aligned}\mathrm{Var}[X(t)]&=E\left[(1-t)B\left(\frac{t}{1-t}\right)(1-s)B\left(\frac{s}{1-s}\right)\right]\\&=(1-t)(1-s)\min\{\frac{t}{1-t},\frac{s}{1-s}\}=s(1-t),(s\leqslant t).\end{aligned}$$

所以 $\{X(t)\}$ 是 Gauss 过程，均值为零，协方差为 $s(1-t)$，即为布朗桥.

习题 7

1. (1) $E[Y(t)]=E[X(t+1)-X(t)]$

$= E[X(t+1)] - E[X(t)]$

$= [\alpha + \beta(t+1)] - (\alpha + \beta t) = \beta$

结果与 t 无关.

(2)$R_Y(t+\tau,t) = E[X(t+\tau+1) - X(t+\tau)] \cdot [\overline{X(t+1) - X(t)}]$

$= R_X(t+\tau+1,t+1) - R_X(t+\tau,t+1) - R_X(t+\tau+1,t) + R_X(t+\tau,t)$

又

$R_X(t_1,t_2) = C_X(t_1,t_2) + m_X(t_1)m_X(t_2)$

$= e^{-\lambda|t_1-t_2|} - (\alpha+\beta t_1)(\alpha+\beta t_2)$,

因此,

$R_Y(t+\tau,t) = 2e^{-\lambda|\tau|} - e^{-\lambda|\tau-1|} - e^{-\lambda|\tau+1|} + \beta^2 = R_Y(\tau)$,结果与 t 无关.

(3)$E[|Y(t)|^2] = R_Y(0) = 2 - 2e^{-\lambda} + \beta^2 < +\infty$.

由(1)、(2)、(3) 知 $Y(t)$ 为平稳过程.

2. $E[Y(n)\overline{Y(m)}] = E[\sum\limits_{k=1}^{n}\sum\limits_{i=1}^{m} a_k \overline{a_i} X(k)\overline{X(i)}]$

$= \sum\limits_{k=1}^{n}\sum\limits_{i=1}^{m} a_k \overline{a_i} E[X(k)\overline{X(i)}]$

$= \sum\limits_{k=1}^{n}\sum\limits_{i=1}^{m} a_k \overline{a_i} R(k,i)$.

由洛弗准则知,如果 $Y(n)$ 均方收敛,应有

$$\lim_{n\to+\infty} E[Y(n)\overline{Y(m)}] = C(\text{常数})$$

即应有级数 $\sum\limits_{k=1}^{n}\sum\limits_{i=1}^{m} a_k \overline{a_i} R(k,i) = c < +\infty$,其中 c 为常数.

6. 由均方变上限积分的性质知

$$m_X(t) - m_X(a) = \int_a^t m_{X'}(s)\mathrm{d}s$$

$$\frac{\mathrm{d}m_X(t)}{\mathrm{d}t} = m_{X'}(t)$$

因此,

$C_{X'}(s,t) = E[X'(s) - m_{X'}(s)][X'(t) - m_{X'}(t)]$

$= E[X'(s)X'(t)] - m_{X'}(s)m_{X'}(t)$

$= R_{X'}(s,t) - m_{X'}(s)m_{X'}(t)$

$= \frac{\partial^2}{\partial s\partial t}[R_X(s,t) - m_X(s)m_X(t)]$

$= \frac{\partial^2}{\partial s\partial t}C_X(s,t)$.

参 考 文 献

[1] 李贤平. 概率论基础. 北京:高等教育出版社,1997.

[2] 王梓坤. 随机过程论. 北京:科学出版社,1965.

[3] 方兆本,缪柏其. 随机过程. 北京: 科学出版社,2011.

[4] 严士健,等. 概率论基础. 北京:科学出版社,2009.

[5] 钱敏平,龚光鲁. 随机过程论. 北京:北京大学出版社,1997.

[6] 林元烈. 应用随机过程. 北京:清华大学出版社,2002.

[7] 刘次华. 随机过程. 武汉:华中科技大学出版社,2008.

[8] 张波. 应用随机过程. 北京:清华大学出版社,2004.

[9] 应坚刚,等. 随机过程基础. 上海:复旦大学出版社,2005.

[10] S. M. Ross. Stochastic Process. New York: Wiley, 1993.

[11] 樊平毅. 随机过程理论与应用. 北京:清华大学出版社,2005.

[12] 胡迪鹤. 随机过程论 基础·理论·应用. 武汉:武汉大学出版社,2005.

[13]复旦大学. 概率论(第三册随机过程). 北京:人民教育出版社,1981.

[General Information]
□□=□□□□□□
□□=□□□□□□□□□□□□□
□□=191
SS□=13638335
DX□=
□□□□=2014.09
□□□=□□□□□□□□□